

Database : GenBank						
No.	Score	Query Match Length	DB	ID	Description	
1	130	100.0	452	6	E64160	E64160 Chemically
2	130	100.0	452	6	AR447842	AR447842 Sequence
3	89	68.5	89	6	E64172	E64172 Chemically
4	89	68.5	89	6	AR447850	AR447850 Sequence
5	63.6	48.9	9937	12	AY560325	AY560325 Cloning
6	63.6	48.9	9946	12	AY560326	AY560326 Cloning
7	62	47.7	735	12	AY833361	AY183361 Synthetic
8	62	47.7	827	12	AY92160	AY192160 Synthetic
9	62	47.7	15382	12	AY178047	AY178047 Expression
10	62	47.7	16368	12	AY178049	AY178049 Expression
11	62	47.7	16520	12	AY178048	AY178048 Expression
12	61	46.9	1006	6	A94910	A94910 Sequence
13	61	46.9	1006	6	BD114819	BD114819 Constitutive
14	60.4	46.5	1138	6	A5980	A59870 Sequence
15	60.4	46.5	1138	6	A94795	A94795 Sequence
16	60	46.5	12607	12	AY737283	AY737283 C-terminal
17	58.2	44.8	2200	6	AY451740	AY451740 Sequence
18	56.2	43.2	5033	6	I28266	I28266 Sequence
19	52	40.0	275	14	TOMMV6	X00052 Tobacco motif

No. is the number of results predicted by chance to have a prep. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	130	100.0	452	6 E641760	B64760 Chemically AR127842 Sequence
2	130	100.0	452	6 AR427842	Sequence
3	89	68.5	89	6 E64172	E64772 Chemically AR127850 Sequence
4	89	68.5	89	6 AR427850	AY560325 Cloning
5	63.6	48.9	9937	12 AY560325	AY560325 Cloning
6	63.6	48.9	9946	12 AY560326	AY183361 Synthetic
7	62	47.7	735	12 AY83361	AY192160 Synthetic
8	62	47.7	827	12 AY92160	AY178047 Expression
9	62	47.7	15382	12 AY178047	AY178049 Expression
10	62	47.7	16368	12 AY178049	AY178048 Expression
11	62	47.7	16520	12 AY178048	A9910 Sequence 22
12	61	46.9	1006	6 A9910	BD134819 Constitutive
13	61	46.9	1006	6 BD134819	A59870 Sequence 2
14	60.4	46.5	1138	6 A59870	A9910 Sequence 39
15	60.4	46.5	1138	6 A94195	AY737233 C-terminal
16	60	46.2	12607	12 AY737283	AY451740 Sequence
17	58.2	44.8	2200	6 AX451740	I28266 Sequence 1
18	56.2	43.2	5033	6 AX451740	X00052 Tobacco mo-
19	52	40.0	225	14 TOMTIV6	

52	0	40.0	6395	6	AR173320	Sequence
52	1	40.0	6395	6	AR271575	Sequence
52	2	40.0	6395	6	AX040174	Sequence
52	3	40.0	6395	6	AX098414	Sequence
52	4	40.0	6395	6	AX193392	Sequence
52	5	40.0	6395	14	AF165190	Tobacco
52	6	40.0	6395	14	AF23221	Tobacco
52	7	40.0	6395	14	AF35127	Tobacco
52	8	40.0	6395	14	AF393128	Tobacco
52	9	40.0	6395	14	AF395129	Tobacco
52	0	40.0	6395	14	AJ01933	Tobacco
52	1	40.0	6395	14	X68110	Tobacco
52	2	40.0	6395	14	TOMV4	Sequence
52	3	40.0	6425	6	AR173322	Sequence
52	4	40.0	6425	6	AX098416	Sequence
52	5	40.0	6439	6	AR173321	Sequence
52	6	40.0	6439	6	AX098415	Sequence
52	7	40.0	6446	6	AR173324	Sequence
52	8	40.0	6446	6	AX098418	Sequence
52	9	40.0	6475	6	AR173323	Sequence
52	0	40.0	6475	6	AX098417	Sequence
52	1	40.0	7926	6	BD235779	Method fo
52	2	40.0	7926	6	AR222869	Sequence
52	3	40.0	7926	6	AR39561	Sequence
52	4	40.0	11641	6	AX465963	Sequence
52	5	40.0	1825	6	AR042908	Sequence

ALIGNMENTS

RESULT 1		PAT 31-JAN-2004
E64760	E64760	452 bp DNA linear
LOCUS	Chemically synthesized artificial promoter for realizing high-level expression of introduced gene and method for synthesizing the same.	PATENT
DEFINITION		
ACCESSION	E64760	
VERSION	E64760-1	GT:18628523
KEYWORDS	JP 2000139477-A/1.	
SOURCE	synthetic construct	
ORGANISM	synthetic construct	
REFERENCE	other sequences; artificial sequences. 1 (bases 1 to 452)	
AUTHORS	Tori,R., Sawanto,S.B., Synge,P.K. and Gupta,S.K.	
TITLE	Chemically synthesized artificial promoter for realizing high-level expression of introduced gene and method for synthesizing the same.	
JOURNAL	COUNCIL OF SCIENTIFIC AND INDUSTRIAL RESEARCH	
COMMENT	OS Artificial Sequence JP 2000139477-A/1 PD 23-MAY-2000 PF 27-APR-1999 JP 1999119227 PR 03-NOV-1998 IN 33/27/98 PI RAKESH TORI, SALLY BISHTUWANATO SAWANTO, PURAJUNNA KUMAR SYNGE, PI SHIFU KUMAR GUPTA PC C12N15/09, C12N5/10, C12N15/00, C12N5/00 CC FH FT Key source FT FEATURES source	
	Location/Qualifiers 1 . 452 FT /organism="Artificial Sequence".	
	Location/Qualifiers 1 . 452 FT /organism="synthetic construct" /mol_type="genomic DNA" /db_xref="taxon:32630"	
ORIGIN	Query Match 100.0%; Score 130; DB 6; Length 452; Best Local Similarity 100.0%; Pred. No. 6.9e-20; Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
Ov	1 TCACATATAATAGGAAGTCATTCAATTGGATGGACAGTSTGATTTCCTCAACA	

RESULT 9		AY178047/c			
Dy	12 AGGAGTCATTGATTTGGATGCCACCTGTGTCATTTCACAATTACCAAC 71	AY178047	15382 bp	DNA	circular SYN 06-APR-2003
Dy	731 AGGAGTCATTGATTTGGATGCCACCTGTGTCATTTCACAATTACCAAC 785	Expression vector	pYPX143,	complete sequence.	
Db		AY178047			
Dy	72 AACAAACACAAACACATTATACTTAACTTACAATTACA 115	AY178047.1	GI:29568870		
Db	786 AACAAACACAAACACAT--TACATTACTTACAATTACA 827				
DEFINITION					
LOCUS AY178047					
DEFINITION Expression vector pYPX143, complete sequence.					
ACCESSION AY178047					
VERSION AY178047.1					
KEYWORDS					
ORGANISM Expression vector pYPX143					
SOURCE Expression vector pYPX143					
ARTIFICAL SEQUENCES ; other sequences; artificial sequences; vectors.					
SEQUENCE 1 (bases 1 to 15382)					
REFERENCE Yao,Q., Peng,R. and Xiong,A.					
AUTHORS					
TITLE Submitted (12-NOV-2002) Shanghai Academy of Agricultural Sciences					
JOURNAL Agro-Biotechnology Research Center, 2901 Beidi Rd, Shanghai 201106, China					
COMMENT NCBI staff are still waiting for submitters to provide appropriate coding region information.					
FEATURES Location/Qualifiers					
source 1..15382					
<i>organism</i> = "Expression vector pYPX143"					
<i>mol type</i> = "genomic DNA"					
<i>db_xref</i> = "Axon:218199"					
<i>lab_host</i> = "Agrobacterium tumefaciens"					
<i>complement</i> (5139..5798)					
<i>note</i> = "synthetic sequence derived from castorbean"					
<i>codon_start</i> =1					
<i>product</i> = "catalase"					
<i>protein_id</i> = "AA084026_1"					
<i>db_xref</i> = "GT: 29568871"					
<i>translation</i> = "MEEKKITGTTVDISQWHKFFRAFQSVAQCTYNQTVOLDITAFLKTWTKKHKRYPAFHILALIIMPAPEPLAMAKDGLIVIWSVHPCYTVEHEQTFLSSLWSEYHDQFQLLHSQDVACYGENLAYPEKGFLENMFPSANPVSFSFTDLNVANMDNDFAPVPMGKYTQGDVKMLPAIAQVHAVCDGFHVSRLMLNLQYCDEWQGG					
misc_feature A"					
<i>complement</i> (6529..7504)					
<i>note</i> = "similar to aminoglycoside-3'-O-phosphotransferase"					
<i>codon_start</i> =1					
<i>product</i> = "Gus"					
<i>protein_id</i> = "AA084027_1"					
<i>db_xref</i> = "GT: 29568872"					
<i>translation</i> = "MLRPVETPTREIKKDGFLWAFSLDRENGIDQRWESALQESRA					
TAVPGSFNDQFADADTRVYAGRNWYQEVFLPGWAGQTRVLRVWQDQVTRVGRWQ					
EVMEHGQGYTPPEADTVPYLAGKSRTVPTCNYNELLWQTMQTPGMVDTENRKQSY					
FHDFHENYAGIHRSSVMLYTPPTWVDDITTVYTAQODCNHAASWDWQVYANGDVSVELRD					
ADQQVVAATGQQTSGLQVNPILWQPGEGGLVLYELVKAQDLSQVYDQVYDQVYDQVYDQVY					
KGEOFQJLHKFYFTGRGRHEDADLRKGFDQDNVYDHALMDWIGANSYRTSHYQPVY					
EEMLDWAEHGLVVIDETAAVGNFLS1GIGPERAGNEPKLEYSERAVNGETQOAHLOQAI					
AHTDTISLDLFDVTLCLARYQGTYVQSDLETAKYLECELLAQOKLHOPLLTTEYGVDF					
TLAGHSMYTDMMWSEXYOCANLDMYMRVEDRVSAYVGEQWNPADFATSGQ1LIRVGGN					
KKGIFTFRDKRPSAFLQKREWGMNGEKPOQGGKQ"					
<i>codon_start</i> =1					
<i>product</i> = "his tagged GusA"					
<i>protein_id</i> = "AA084028_1"					
<i>db_xref</i> = "GT: 29568873"					
<i>translation</i> = "MWDLNRNRLRVPVTPREIKKDGFLWAFSLDRENGIDQRWESALQESRA					
SALQESRAIAVPGSSENDQFADADTRVYAGRNWYQEVFLPGWAGQTRVLRVWQDQVTRVGRWQ					
GKWKWNNQEINMEHGQGYTPPEADTVPYLAGKSRTVPTCNYNELLWQTMQTPGMVDTENRKQSY					
NGKQKSYFHDHENYAGIHRSSVMLYTPPTWVDDITTVYTAQODCNHAASWDWQVYANGDVSVELRD					
TLAGHSMYTDMMWSEXYOCANLDMYMRVEDRVSAYVGEQWNPADFATSGQ1LIRVGGN					
KKGIFTFRDKRPSAFLQKREWGMNGEKPOQGGKQ"					
<i>codon_start</i> =1					
<i>product</i> = "his tagged GusA"					
<i>protein_id</i> = "AA084028_1"					
<i>db_xref</i> = "GT: 29568873"					
<i>translation</i> = "MWDLNRNRLRVPVTPREIKKDGFLWAFSLDRENGIDQRWESALQESRA					
SALQESRAIAVPGSSENDQFADADTRVYAGRNWYQEVFLPGWAGQTRVLRVWQDQVTRVGRWQ					
GKWKWNNQEINMEHGQGYTPPEADTVPYLAGKSRTVPTCNYNELLWQTMQTPGMVDTENRKQSY					
NGKQKSYFHDHENYAGIHRSSVMLYTPPTWVDDITTVYTAQODCNHAASWDWQVYANGDVSVELRD					
TLAGHSMYTDMMWSEXYOCANLDMYMRVEDRVSAYVGEQWNPADFATSGQ1LIRVGGN					
KKGIFTFRDKRPSAFLQKREWGMNGEKPOQGGKQ"					
<i>codon_start</i> =1					
<i>product</i> = "his tagged GusA"					
<i>protein_id</i> = "AA084028_1"					
<i>db_xref</i> = "GT: 29568873"					
<i>translation</i> = "MWDLNRNRLRVPVTPREIKKDGFLWAFSLDRENGIDQRWESALQESRA					
SALQESRAIAVPGSSENDQFADADTRVYAGRNWYQEVFLPGWAGQTRVLRVWQDQVTRVGRWQ					
GKWKWNNQEINMEHGQGYTPPEADTVPYLAGKSRTVPTCNYNELLWQTMQTPGMVDTENRKQSY					
NGKQKSYFHDHENYAGIHRSSVMLYTPPTWVDDITTVYTAQODCNHAASWDWQVYANGDVSVELRD					
TLAGHSMYTDMMWSEXYOCANLDMYMRVEDRVSAYVGEQWNPADFATSGQ1LIRVGGN					
KKGIFTFRDKRPSAFLQKREWGMNGEKPOQGGKQ"					
<i>codon_start</i> =1					
<i>product</i> = "his tagged GusA"					
<i>protein_id</i> = "AA084028_1"					
<i>db_xref</i> = "GT: 29568873"					
<i>translation</i> = "MWDLNRNRLRVPVTPREIKKDGFLWAFSLDRENGIDQRWESALQESRA					
SALQESRAIAVPGSSENDQFADADTRVYAGRNWYQEVFLPGWAGQTRVLRVWQDQVTRVGRWQ					
GKWKWNNQEINMEHGQGYTPPEADTVPYLAGKSRTVPTCNYNELLWQTMQTPGMVDTENRKQSY					
NGKQKSYFHDHENYAGIHRSSVMLYTPPTWVDDITTVYTAQODCNHAASWDWQVYANGDVSVELRD					
TLAGHSMYTDMMWSEXYOCANLDMYMRVEDRVSAYVGEQWNPADFATSGQ1LIRVGGN					
KKGIFTFRDKRPSAFLQKREWGMNGEKPOQGGKQ"					
<i>codon_start</i> =1					
<i>product</i> = "his tagged GusA"					
<i>protein_id</i> = "AA084028_1"					
<i>db_xref</i> = "GT: 29568873"					
<i>translation</i> = "MWDLNRNRLRVPVTPREIKKDGFLWAFSLDRENGIDQRWESALQESRA					
SALQESRAIAVPGSSENDQFADADTRVYAGRNWYQEVFLPGWAGQTRVLRVWQDQVTRVGRWQ					
GKWKWNNQEINMEHGQGYTPPEADTVPYLAGKSRTVPTCNYNELLWQTMQTPGMVDTENRKQSY					
NGKQKSYFHDHENYAGIHRSSVMLYTPPTWVDDITTVYTAQODCNHAASWDWQVYANGDVSVELRD					
TLAGHSMYTDMMWSEXYOCANLDMYMRVEDRVSAYVGEQWNPADFATSGQ1LIRVGGN					
KKGIFTFRDKRPSAFLQKREWGMNGEKPOQGGKQ"					
<i>codon_start</i> =1					
<i>product</i> = "his tagged GusA"					
<i>protein_id</i> = "AA084028_1"					
<i>db_xref</i> = "GT: 29568873"					
<i>translation</i> = "MWDLNRNRLRVPVTPREIKKDGFLWAFSLDRENGIDQRWESALQESRA					
SALQESRAIAVPGSSENDQFADADTRVYAGRNWYQEVFLPGWAGQTRVLRVWQDQVTRVGRWQ					
GKWKWNNQEINMEHGQGYTPPEADTVPYLAGKSRTVPTCNYNELLWQTMQTPGMVDTENRKQSY					
NGKQKSYFHDHENYAGIHRSSVMLYTPPTWVDDITTVYTAQODCNHAASWDWQVYANGDVSVELRD					
TLAGHSMYTDMMWSEXYOCANLDMYMRVEDRVSAYVGEQWNPADFATSGQ1LIRVGGN					
KKGIFTFRDKRPSAFLQKREWGMNGEKPOQGGKQ"					
<i>codon_start</i> =1					
<i>product</i> = "his tagged GusA"					
<i>protein_id</i> = "AA084028_1"					
<i>db_xref</i> = "GT: 29568873"					
<i>translation</i> = "MWDLNRNRLRVPVTPREIKKDGFLWAFSLDRENGIDQRWESALQESRA					
SALQESRAIAVPGSSENDQFADADTRVYAGRNWYQEVFLPGWAGQTRVLRVWQDQVTRVGRWQ					
GKWKWNNQEINMEHGQGYTPPEADTVPYLAGKSRTVPTCNYNELLWQTMQTPGMVDTENRKQSY					
NGKQKSYFHDHENYAGIHRSSVMLYTPPTWVDDITTVYTAQODCNHAASWDWQVYANGDVSVELRD					
TLAGHSMYTDMMWSEXYOCANLDMYMRVEDRVSAYVGEQWNPADFATSGQ1LIRVGGN					
KKGIFTFRDKRPSAFLQKREWGMNGEKPOQGGKQ"					
<i>codon_start</i> =1					
<i>product</i> = "his tagged GusA"					
<i>protein_id</i> = "AA084028_1"					
<i>db_xref</i> = "GT: 29568873"					
<i>translation</i> = "MWDLNRNRLRVPVTPREIKKDGFLWAFSLDRENGIDQRWESALQESRA					
SALQESRAIAVPGSSENDQFADADTRVYAGRNWYQEVFLPGWAGQTRVLRVWQDQVTRVGRWQ					
GKWKWNNQEINMEHGQGYTPPEADTVPYLAGKSRTVPTCNYNELLWQTMQTPGMVDTENRKQSY					
NGKQKSYFHDHENYAGIHRSSVMLYTPPTWVDDITTVYTAQODCNHAASWDWQVYANGDVSVELRD					
TLAGHSMYTDMMWSEXYOCANLDMYMRVEDRVSAYVGEQWNPADFATSGQ1LIRVGGN					
KKGIFTFRDKRPSAFLQKREWGMNGEKPOQGGKQ"					
<i>codon_start</i> =1					
<i>product</i> = "his tagged GusA"					
<i>protein_id</i> = "AA084028_1"					
<i>db_xref</i> = "GT: 29568873"					
<i>translation</i> = "MWDLNRNRLRVPVTPREIKKDGFLWAFSLDRENGIDQRWESALQESRA					
SALQESRAIAVPGSSENDQFADADTRVYAGRNWYQEVFLPGWAGQTRVLRVWQDQVTRVGRWQ					
GKWKWNNQEINMEHGQGYTPPEADTVPYLAGKSRTVPTCNYNELLWQTMQTPGMVDTENRKQSY					
NGKQKSYFHDHENYAGIHRSSVMLYTPPTWVDDITTVYTAQODCNHAASWDWQVYANGDVSVELRD					
TLAGHSMYTDMMWSEXYOCANLDMYMRVEDRVSAYVGEQWNPADFATSGQ1LIRVGGN					
KKGIFTFRDKRPSAFLQKREWGMNGEKPOQGGKQ"					
<i>codon_start</i> =1					

ADQVYATGCGTSGTILQVNPHLWQFEGGYLYELCVTAKSOTECDIYPLRVGIRSVAV
RGEQFLINHKFYEFTCFGRHDDAQLRGKGPENVLMLVHDHALMDWGANSTRTSVYPA
EMLWDADERGLIVVIDAEGKPELQVNPFLQGAGTTECIPYELVGIRSTAV
KELIARDKHNPSVVMNSIANPDTDREQARYFAPIAETAKLDPRTPCTCVNTMFCD
AHTDTISDLFVCLNRYGGYVQSSDLETEKEVLEKELLAWOKEWLHOPIIITEGVGD
TLAGLSMYTSDVSEYQVWMSIANBPDTRPCTCVNTMFCD
RKGIFTRDRKPSKAFLQKWTGMFKEQKQOGQQ
14114 .1561

/note="similar to his tagged GusA"

misc_feature

ORIGIN Query Match 47.7%; Score 62; DB 12; Length 16368;
Best Local Similarity 88.5%; Prod. No. 0.00012; Mismatches 0; Indels 5; Gaps 2;
Matches 92; Conservative 0; Mismatches 5; Indels 7; Gaps 2;

Qy 112 AGGAGTCAATTGATGGAAATGGACAGCTGTGTCAATTCTAACATTACCAACAC 71
Db 11351 AGGAGTCAATTGATGGAGGACGCTAT---TTTACACATTACCAACAC 11297

Qy 72 AACRAACACAAACAAACATTATAACATTACTATTACATTACA 115
Db 11296 AACAAACACAAACAAACAT--TACAATTACTATTACATTACA 11255

RESULT 11
AY178048/c LOCUS AY178048 16520 bp DNA circular SYN 06-APR-2003

DEFINITION Expression vector pYPX145, complete sequence.

VERSION AY178048
KEYWORDS AY178048.1 GI:29568874

SOURCE ORGANISM Expression vector pYPX145
Expression vector pYPX145
other sequences; artificial sequences; vectors.

REFERENCE AUTHORS Yao,Q., Peng,R. and Xiong,A.

TITLE Direct Submission
JOURNAL Submitted (12-NOV-2002) Shanghai Academy of Agricultural Sciences,
Agro-Biotechnology Research Center, 2901 Beidi Rd, Shanghai 201016,
China

COMMENT NCBI staff are still waiting for submitters to provide appropriate
coding region information.

FEATURES Location/Qualifiers

1. 16520 /organism="Expression vector pYPX145"
(mol_type="genomic DNA"
/db_xref="Taxon:218200"
/lab_host="Agrobacterium tumefaciens"
compлемент (5139 .5798)
/gene="cat",
complement (5139 .5798)

/codon_start=1
/product="chloramphenicol acetyl transferase"

/protein_id="AA084028.1"
/db_xref="GI:29568874"

/translation="MEKKITCCTTYTDISOMHRKEHFEAFOSQVAQCTNYTQYQDITAF
IUKTKGNKHKKYPAFHILANMNAHBFNAMKDGELVITWSDVRCYTFHEQIBTF
SSLWSIYHDDPROFIYQSOVACYGENLAYFPKFIEEMNFVSANPWFSTSPDLNV
ANMDNEFAPVPTMGKYTOGKVMLPLAIQTHAVCDGFHVGRMLNELQOYCDERQGG
A"

complement (6529 .7504)
/note="similar to amino-glycoside-3'-O-phosphotransferase"

/complement (9589 .11400)
/Gene="gus",
complement (9589 .11400)

/Gene="gus",
/codon_start=1
/product="gus"

/protein_id="AA084028.1"

/db_xref="GI:29568874"

/translation="MLRPVEPTREIKKDLWAFSLDRENCGIDQRWNESESQESRA
IAVPGSFNDQPADADIRNYAGNWYOREVFPKGMAGQRLVLPDAVTHYCKWVNNQ

misc_feature

ORIGIN Query Match 47.7%; Score 62; DB 12; Length 16520;
Best Local Similarity 88.5%; Prod. No. 0.00012; Mismatches 0; Indels 5; Gaps 2;

Qy 12 AGGAAGTCATTGATGGATGGACAGCTGTGTCAATTACCAACAC 71
Db 11503 AGGAAGTCATTGAGGAGCACGTAT---TTTACACATTACCAACAC 11449

Qy 72 AACAAACACAAACAAACATTATAACATTACTATTACATTACA 115
Db 11448 AACAAACACAAACAAACAT--TACAAATTACTATTACATTACA 11407

RESULT 12
A94910 LOCUS A94910 Sequence 22 from Patent WO931258.

DEFINITION ACCESSION A94910
VERSION A94910.1 GI:6779107

KEYWORDS SOURCE unidentified
ORGANISM unidentified
unclassified.

REFERENCE 1 (bases 1 to 1006)
AUTHORS Stuiver, M.H. and Sibbolts, F.H.

JOURNAL CONSTITUTIVE PLANT PROMOTERS
PATENT : WO 931258-A 22 24-JUN-1999 ;
STUIVER MAARTEN HENDRIK (NL) ; MOGEN INT (NL)

FEATURES FEATURES
Source
1 . 1006
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32654"

ORIGIN Query Match 46.9%; Score 61; DB 6; Length 1006;
Best Local Similarity 88.3%; Prod. No. 0.00012; Mismatches 0; Indels 7; Gaps 2;

Qy 12 AGGAAGTCATTGATGGATGGACAGCTGTGTCAATTACCAACAC 71
Db 906 AGGAAGTCATTGAGGAGCACGTAT---TTTACACATTACCAACAC 960

Qy 72 AACAAACACAAACAAACATTATAACATTACTATTACATTACA 114
Db 961 AACAAACACAAACAAACAT--TACAAATTACTATTACATTACA 1001

RESULT 13
BD134819 LOCUS BD134819

DEFINITION ACCESSION BD134819
VERSION BD134819
KEYWORDS SOURCE unidentified
ORGANISM unidentified
unclassified.

REFERENCE 1 (bases 1 to 1006)

AUTHORS Stuiver, M.H. and Sibbolz, F.H.

TITLE Constitutive plant promoter
JOURNAL Patent: JP 2002308186-A 22-19-MAR-2002;
COMMENT MOGEN INTERNATIONAL NV
OS Unidentified
PN JP 2002308186-A/22
PD 19-MAR-2002
PR 10-DEC-1998 JP 20005319156
PF 12-DEC-1997 EP 97203912-7
PI MAARTEN HENDRIK STUIVER, FLOOR HENDRIC SIBORTZ PC
C12N15/09, C12N15/00
CC Strandedness: Double;
CC Topology: Linear;
CC Constitutive plant promoter
FH Key
FT source
FEATURES Location/Qualifiers
source 1..1006
/organism='Unidentified'.
FEATURES Location/Qualifiers
source 1..1006
/organism='Unidentified'
/mol_type='genomic DNA'
/db_xref="taxon:32644"
ORIGIN

Query Match 46.9%; Score 61; DB 6; Length 1006;
 Best Local Similarity 88.3%; Pred. No. 0_00031; Mismatches 0; Indels 5; Gaps 2;
 Matches 91; Conservative 0; Unconservative 91;

Qy 1..12 AGGAGTTCATTTCATTTGAAATTGGACAGCTGTGTCAATTCTCAACAAATTACCAACAC 71
 Db 906 AGGAGTTCATTTCATTTGAGGAGCTAT-----TTTACAACAAATTACCAACAC 960

Qy 72 AACAAACAAACAAACAAACTATACATTACTATTACATTAC 114
 Db 961 AACAAACAAACAAACAACTT-TACAATTACTATTACATTAC 1001..

RESULT 14
A59870 A59870 1138 bp DNA linear PAT 06-MAR-1998
LOCUS Sequence 2 from Patent WO9706269.
DEFINITION A59870
ACCESSION A59870
VERSION GI:3715061
KEYWORDS unidentified
SOURCE unidentified
ORGANISM unidentified
REFERENCE
AUTHORS JEPSON, I.
TITLE INDUCIBLE HERBICIDE RESISTANCE
JOURNAL Patent: WO 9706269-A 2 20-FEB-1997;
FEATURES Location/Qualifiers
source 1..1138
/organism='Unidentified'
/mol_type='unassigned DNA'
/isolate="PLASMID MJ31"
/db_xref="taxon:32644"
ORIGIN

Query Match 46.5%; Score 60.4; DB 6; Length 1138;
 Best Local Similarity 85.8%; Pred. No. 0_00041; Mismatches 0; Indels 11; Gaps 2;
 Matches 91; Conservative 91;

Qy 1..12 AGGAGTTCATTTCATTTGAAATTGGACAGCTGTGTCAATTCTCAACAAATTACCAACAC 71
 Db 743 AGGAGTTCATTTCATTTGAGGAGG--ACCTCGATTATTACAACAAATTACCAACAC 800

Qy 72 AACAAACAAACAAACAAACTATACATTACTATTACATTAC 117
 Db 801 AACAAACAAACAAACAACTT-TACAATTACTATTACATTAC 844

Search completed: July 3, 2005, 03:19:09
 Job time : 724.588 secs

the rugo brasil (uspi)

RESULT 2
US-09-263-692A-5
; Sequence 5, Application US/09263692A
; Patent No. 6639065
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A chemically synthesized artificial promoter for high level expression of sequences and a method for its synthesis
; FILE REFERENCE: Q52511
; CURRENT APPLICATION NUMBER: US/09/263, 692A
; CURRENT FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: 3322/De1/98
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 46
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: minimal domain (b)
US-09-263-692A-5

Query Match 14.7% Score 46; DB 4; Length 46;
Best Local Similarity 100%; Pred. No. 6.6e-06; Indels 0; Gaps 0;

Qy 267 CCACTGACCACATTTGAGCCACATGAGCCACTTGAAGCTACT 312
Db 1 CCACTGACCACATTTGAGCCACATGAGCCACATTTGAGCTACT 46

RESULT 3
US-09-263-692A-6
; Sequence 6, Application US/09263692A
; Patent No. 6639065
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A chemically synthesized artificial promoter for high level expression of sequences and a method for its synthesis
; FILE REFERENCE: Q52511
; CURRENT APPLICATION NUMBER: US/09/263, 692A
; CURRENT FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: 3322/De1/98
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 46
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: domain I
US-09-263-692A-6

Query Match 14.7% Score 46; DB 4; Length 46;
Best Local Similarity 100%; Pred. No. 6.6e-06; Indels 0; Gaps 0;

Qy 221 GCTTGTAACGTGAGCTGAGCTGAGTAGATAACGTCACGGT 266
Db 1 GCTTGTAACGTGAGCTGAGCTGAGTAGATAACGTCACGGT 46

RESULT 4
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.

Query Match 14.7% Score 46; DB 4; Length 46;
Best Local Similarity 100%; Pred. No. 6.6e-06; Indels 0; Gaps 0;

Qy 221 GCTTGTAACGTGAGCTGAGCTGAGTAGATAACGTCACGGT 266
Db 1 GCTTGTAACGTGAGCTGAGTAGATAACGTCACGGT 46

RESULT 5
US-09-949-016-12417
; Sequence 12417, Application US/09549016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

GENERAL INFORMATION:
 APPLICANT: VENTER, J. Craig et al.
 TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 FILE REFERENCE: CLO01307
 CURRENT APPLICATION NUMBER: US/09/949,016
 CURRENT FILING DATE: 2000-04-14
 PRIOR APPLICATION NUMBER: 60/241,755
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/237,768
 PRIOR FILING DATE: 2000-10-03
 PRIOR APPLICATION NUMBER: 60/231,498
 PRIOR FILING DATE: 2000-09-08
 NUMBER OF SEQ ID NOS: 207012
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 14020
 LENGTH: 22294
 TYPE: DNA
 ORGANISM: Human
 US-09-949-016-14020

RESULT 10
 Query Match 10.8%; Score 33.8; DB 4; Length 22294;
 Best Local Similarity 53.4%; Prod. No. 1.8;
 Matches 71; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
 Db 4766 ATCTGACCATCTCTGATCTGGAAAAAGTTCACTTACCTACATATGTGGAT 182
 Qy 123 ATCTGACCATCTCTGATCTGGATCTGGCTTACCTACATATGTGGAT 182
 Db 4766 AACTGCTATTGTCATTGCTCTTAAATGCCAAATTCTATGATTACATGGAAA 4707
 Qy 183 TGTGGAAGAAGAACGAGGCACTGGTGGAAAGANGCTGTACGTGTACGTGAGC 242
 Db 4706 TGACGAAATAAAACAGTAATGATGGAATAATGATGATCATTTAAAGCGGTG 4647
 Qy 243 ATAGATAGATACA 255
 Db 4646 ATATAAATTAAA 4634

US-09-007-005-32
 Sequence 32, Application US/09007005B
 Patent No. 6258558
 GENERAL INFORMATION:
 APPLICANT: Szcotak, Jack W.
 APPLICANT: Roberts, Richard W.
 APPLICANT: Liu, Rhei
 TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
 FILE REFERENCE: 00786/350003
 CURRENT APPLICATION NUMBER: US/09/007,005B
 CURRENT FILING DATE: 1998-01-14
 EARLIER APPLICATION NUMBER: 60/035,963
 EARLIER FILING DATE: 1997-01-27
 EARLIER APPLICATION NUMBER: 60/064,491
 EARLIER FILING DATE: 1997-11-06
 NUMBER OF SEQ ID NOS: 33
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 32
 LENGTH: 248
 TYPE: RNA
 ORGANISM: Homo sapiens
 US-09-007-005-32

RESULT 11
 Query Match 10.8%; Score 33.6; DB 3; Length 248;
 Best Local Similarity 18.2%; Prod. No. 0.23;
 Matches 28; Conservative 65; Mismatches 61; Indels 0; Gaps 0;
 Db 136 TAGATCGTGGAAAAAGTTCACTTACCTTACCTATGTGGATCTGGAAAGAA 195
 Qy 136 TAGATCGTGGAAAAAGTTCACTTACCTTACCTATGTGGATCTGGAAAGAA 195
 Db 62 URGARARGRARARCRGRURGRURGRURGRURGRARGRARGRAR 121
 Qy 196 GACGGAGGCATGGTGGAAAAGAAAGTGTACGTGTACGTGTACGTGAGTAGATGATA 255
 Db 122 CRCRURGRCRUGRCRURGRURGRURGRURGRURGRURGRAR 181
 Qy 256 CGTGCACGGTCCACTTGACGCCAAATTGCCA 289
 Db 182 ARCRARCRARCRGRGRARCRGRARCRGR 215

RESULT 12
 Query Match 10.8%; Score 33.6; DB 3; Length 248;
 Best Local Similarity 18.2%; Prod. No. 0.23;
 Matches 28; Conservative 65; Mismatches 61; Indels 0; Gaps 0;
 Db 62 URGARARGRARARCRGRURGRURGRURGRURGRURGRARGRAR 121
 Qy 196 GACGGAGGCATGGTGGAAAAGAAAGTGTACGTGTACCTATGTGGATCTGGAAAGAA 195
 Db 62 URGARARGRARARCRGRURGRURGRURGRURGRURGRURGRARGRAR 121
 Qy 196 GACGGAGGCATGGTGGAAAAGAAAGTGTACGTGTACCTATGTGGATCTGGAAAGAA 195
 Db 122 CRCRURGRCRUGRCRURGRURGRURGRURGRURGRURGRAR 181
 Qy 256 CGTGCACGGTCCACTTGACGCCAAATTGCCA 289
 Db 182 ARCRARCRARCRGRGRARCRGRARCRGR 215

RESULT 13
 Query Match 10.8%; Score 33.6; DB 3; Length 248;
 Best Local Similarity 18.2%; Prod. No. 0.23;
 Matches 28; Conservative 65; Mismatches 61; Indels 0; Gaps 0;
 Db 136 TAGATCGTGGAAAAAGTTCACTTACCTTACCTATGTGGATCTGGAAAGAA 195
 Qy 136 TAGATCGTGGAAAAAGTTCACTTACCTTACCTATGTGGATCTGGAAAGAA 195
 Db 62 URGARARGRARARCRGRURGRURGRURGRURGRURGRARGRAR 121
 Qy 196 GACGGAGGCATGGTGGAAAAGAAAGTGTACGTGTACGTGTACGTGAGTAGATGATA 255
 Db 122 CRCRURGRCRUGRCRURGRURGRURGRURGRURGRURGRAR 181
 Qy 256 CGTGCACGGTCCACTTGACGCCAAATTGCCA 289
 Db 182 ARCRARCRARCRGRGRARCRGRARCRGR 215

RESULT 14
 Query Match 10.8%; Score 33.6; DB 3; Length 248;
 Best Local Similarity 18.2%; Prod. No. 0.23;
 Matches 28; Conservative 65; Mismatches 61; Indels 0; Gaps 0;
 Db 62 URGARARGRARARCRGRURGRURGRURGRURGRURGRARGRAR 121
 Qy 196 GACGGAGGCATGGTGGAAAAGAAAGTGTACGTGTACCTATGTGGATCTGGAAAGAA 195
 Db 62 URGARARGRARARCRGRURGRURGRURGRURGRURGRURGRARGRAR 121
 Qy 196 GACGGAGGCATGGTGGAAAAGAAAGTGTACGTGTACCTATGTGGATCTGGAAAGAA 195
 Db 122 CRCRURGRCRUGRCRURGRURGRURGRURGRURGRURGRAR 181
 Qy 256 CGTGCACGGTCCACTTGACGCCAAATTGCCA 289
 Db 182 ARCRARCRARCRGRGRARCRGRARCRGR 215

ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Translation template
 US-09-007-005-3

Query Match 10.8%; Score 33.6; DB 3; Length 277;
 Best Local Similarity 18.2%; Pred. No. 0.24; *
 Matches 28; Conservative 65; Mismatches 61; Indels 0; Gaps 0;

Qy 136 TAGATCGTGGAAAAGTCACTGTTACGTGCTTAATGATTGTTGGAAAAAGAA 195
 Db 62 URGARARGRARARCRGRARARCRURGRURGRARARCRURGRARARCR 121

Qy 196 GACGGAGGCATCGGTGGAAAAGAGCTTGATTGCTGACGTTAGATAGATACA 255
 Db 122 CRCRURGRURGRURGRURGRARARCRURGRURGRARARCRURGRARAR 181

Qy 256 CGTCACGGCTCCACTTGACGCCAATGACGCA 289
 Db 182 ARCRARARARCRURGRURGRARARCRURGRCR 215

RESULT 12
 US-09-244-796-3
 Sequence 3, Application US/09244796
 P. Patent No. 6281344
 GENERAL INFORMATION:
 APPLICANT: Szostak, Jack W.
 APPLICANT: Roberts, Richard W.
 APPLICANT: Liu, Rite
 TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
 FILE REFERENCE: 00786/350007
 CURRENT APPLICATION NUMBER: US/09/244,796
 CURRENT FILING DATE: 1999-02-05
 EARLIER APPLICATION NUMBER: 60/035,963
 EARLIER FILING DATE: 1997-01-27
 EARLIER APPLICATION NUMBER: 60/064,491
 EARLIER FILING DATE: 1997-11-06
 EARLIER APPLICATION NUMBER: 09/007,005
 EARLIER FILING DATE: 1998-01-14
 NUMBER OF SEQ ID NOS: 33
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 3
 LENGTH: 277
 TYPE: RNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Translation template
 US-09-244-796-3

Query Match 10.8%; Score 33.6; DB 3; Length 277;
 Best Local Similarity 18.2%; Pred. No. 0.24; *
 Matches 28; Conservative 65; Mismatches 61; Indels 0; Gaps 0;

Qy 136 TAGATCGTGGAAAAGTCACTGTTACGTACTATGTTGGATTGTTGGAAAAAGAA 195
 Db 62 URGARARGRARARCRGRARARCRURGRURGRARARCRURGRARARCR 121

Qy 196 GACGGAGGCATCGGTGGAAAAGAGCTTGATTGCTGACGTTAGATAGATACA 255
 Db 122 CRCRURGRURGRURGRARARCRURGRURGRARARCRURGRARAR 181

Qy 256 CGTCACGGCTCCACTTGACGCCAATGACGCA 289
 Db 182 ARCRARARARCRURGRURGRARARCRURGRCR 215

RESULT 13
 US-09-244-78043
 Sequence 78043, Application US/09949016
 P. Patent No. 681339
 GENERAL INFORMATION:

TELEPHONE: 610-270-4478
 TELEX: 610-270-5090
 INFORMATION FOR SEQ ID NO: 136:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1096 base pairs
 TYPE: nucleic acid
 STRANDBNESS: single
 TOPOLOGY: linear
 JS-08-858-207A-136

	Query	Match	Score	Length	DB	3:	Length	1096:
Best Local Similarity	51.4%	Pred. No.	1.2;					
Matches	75;	Mismatches	0;	Indels	0;	Gaps	0;	
Qy	AAAGAACATTCAATCCAGAAAAGGTGGAAAGTTGGTGGATTCGGTGGAAAAGTTTC	62	AAAGAACATTCAATCCAGAAAAGGTGGAAAGTTGGTGGATTCGGTGGAAAAGTTTC	62	Qy	62	AAAGAACATTCAATCCAGAAAAGGTGGAAAGTTGGTGGATTCGGTGGAAAAGTTTC	121
Db	AGACAGGTAATTACTTGAAAGATGGTTAGAGTTATCTGGTGGTAATCAATGAA	6373	AGACAGGTAATTACTTGAAAGATGGTTAGAGTTATCTGGTGGTAATCAATGAA	6373	Db	6373	AGACAGGTAATTACTTGAAAGATGGTTAGAGTTATCTGGTGGTAATCAATGAA	121
Qy	GATCTGACCATCTCATCGATCCTGGAAAAAGTCACTGCGTTACGTTACATATGGAA	1122	GATCTGACCATCTCATCGATCCTGGAAAAAGTCACTGCGTTACGTTACATATGGAA	1122	Qy	1122	GATCTGACCATCTCATCGATCCTGGAAAAAGTCACTGCGTTACGTTACATATGGAA	181
Db	GTTTGAACAATCTTACTTTATGGAATTAAACAGGGAAAGATAAGTTCAAGTGTGTTT	6313	GTTTGAACAATCTTACTTTATGGAATTAAACAGGGAAAGATAAGTTCAAGTGTGTTT	6313	Db	6313	GTTTGAACAATCTTACTTTATGGAATTAAACAGGGAAAGATAAGTTCAAGTGTGTTT	6254
Qy	TGTGGAAAAGAACAGGGGACATC	182	TGTGGAAAAGAACAGGGGACATC	182	Qy	182	TGTGGAAAAGAACAGGGGACATC	207
Db	TTCCTGAAAAAATTGCACCTCCCCATC	6253	TTCCTGAAAAAATTGCACCTCCCCATC	6253	Db	6253	TTCCTGAAAAAATTGCACCTCCCCATC	6228

Search completed: July 3, 2005, 04:55:25
 Job time : 143.471 secs

	Query	Match	Score	Length	DB	3:	Length	1096:
Best Local Similarity	51.4%	Pred. No.	1.2;					
Matches	75;	Mismatches	0;	Indels	0;	Gaps	0;	
Qy	AAAGAACATTCAATCCAGAAAAGGTGGAAAGTTCTGGATTCGGTGGAAAAGTTTC	62	AAAGAACATTCAATCCAGAAAAGGTGGAAAGTTCTGGATTCGGTGGAAAAGTTTC	62	Qy	62	AAAGAACATTCAATCCAGAAAAGGTGGAAAGTTCTGGATTCGGTGGAAAAGTTTC	121
Db	AGAGAGGTAATTACTTGAAAGATGGTTAGAGTTATCTGGTGGTAATCAATGAA	699	AGAGAGGTAATTACTTGAAAGATGGTTAGAGTTATCTGGTGGTAATCAATGAA	699	Db	699	AGAGAGGTAATTACTTGAAAGATGGTTAGAGTTATCTGGTGGTAATCAATGAA	640
Qy	GATCTGACCATCTCATCGATCCTGGAAAAAGTCACTGCGTTACGTTACATATGGAA	122	GATCTGACCATCTCATCGATCCTGGAAAAAGTCACTGCGTTACGTTACATATGGAA	122	Qy	122	GATCTGACCATCTCATCGATCCTGGAAAAAGTCACTGCGTTACGTTACATATGGAA	181
Db	GTTTGAACAATCTTACTTTATGGAATTAAACAGGGAAAGATAAGTTCAAGTGTGTTT	639	GTTTGAACAATCTTACTTTATGGAATTAAACAGGGAAAGATAAGTTCAAGTGTGTTT	639	Db	639	GTTTGAACAATCTTACTTTATGGAATTAAACAGGGAAAGATAAGTTCAAGTGTGTTT	580
Qy	TGTGGAAAAGAACAGGGGACATC	182	TGTGGAAAAGAACAGGGGACATC	182	Qy	182	TGTGGAAAAGAACAGGGGACATC	207
Db	TTCCTGAAAAAATTGCACCTCCCCATC	579	TTCCTGAAAAAATTGCACCTCCCCATC	579	Db	579	TTCCTGAAAAAATTGCACCTCCCCATC	554

RESULT 15
JS-08-961-527-115/C
Sequence 115, Application US/08961527
; Sequence 115, Application US/08961527
; Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 115:
SEQUENCE CHARACTERISTICS:
LENGTH: 11303 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
JS-08-961-527-115/C
Query Match 10.4%; Score 32.4; DB 3; Length 11303;
Best Local Similarity 51.4%; Pred. No. 3.9;
Matches 75; Mismatches 0; Indels 0; Gaps 0;

ALIGNMENTS

RESULT 1
E64760 E64760 452 bp DNA linear PAT 31-JAN-2002
LOCUS DEFINITION Chemically synthesized artificial promoter for realizing high-level expression of introduced gene and method for synthesizing the

No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

IMMATURES

Result No.	Score	Query Match	Length	DB ID	Description
1	310.4	99.5	452	6	E64760 Chemically AR427842 Sequence
2	310.4	99.5	452	6	AR427842 Sequence
3	46	14.7	46	6	B64764 Chemically
4	46	14.7	46	6	E64765 Chemically
5	46	14.7	46	6	AR427845 Sequence
6	46	14.7	46	6	AR427846 Sequence
C	7	41.6	13.3	7218	6
C	8	41	13.1	114519	8 AC124964 Medicago
C	9	41	13.1	114519	8 AC124964 Medicago
C	10	40.6	13.0	172623	2 AC135605 Rattus no
C	11	40.6	13.0	197004	2 AC147540 Pan trogl
C	12	39.2	12.6	134442	8 AC145064 Pan trogl
C	13	39.2	12.6	150309	8 AP004619 Oryza sat
C	14	39.2	12.6	189332	2 AC112060 Rattus no
C	15	39.2	12.6	185612	2 AC108612 Rattus no
C	16	38.8	12.4	330362	2 AC1129990 Rattus no
C	17	36.8	11.8	63114	5 CR377227 Zebrifish
C	18	36.8	11.8	140068	9 AL116241 Human DNA
C	19	36.8	11.8	140068	9 CNS01RHN Human Chr

Db	1 GTCGACCATCATTGAAAGGCCCTCGGTAATACATTGGAAAAAAGTGGATAATACCGGA 60	RESULT 3
Qy	61 AAAGAAGATTCACTCATCCAGAAAAGGTGTGGAAAAAGTGTGGATTGGTGGAAACTT 120	E64764 Locus 46 bp DNA linear PAT 31-JAN-2002
Db	61 AAAGAAGATTCACTCATCCAGAAAAGGTGTGGATTGGTGGAAACTT 120	DEFINITION Chemically synthesized artificial promoter for realizing high-level expression of introduced gene and method for synthesizing the same.
Qy	121 CGATCTGACCATCTAGATCGTGGAAAAGTTCACGGTAGCCCTTACGTACATATGGG 180	ACCESSION E64764
Db	121 CGATCTGACCATCTAGATCGTGGAAAAGTTCACGGTAGCCCTTACGTACATATGGG 180	VERSION E64764 .1 GI:18628527
Qy	181 ATTTGGGAAAGAGAGGAGGATCGTGGAAAAGTTCACGGTAGCCCTTACGTACATATGGG 180	KEYWORDS JP 2000139477-A/5.
Db	181 ATTTGGGAAAGAGAGGAGGATCGTGGAAAAGTTCACGGTAGCCCTTACGTACATATGGG 180	SOURCE synthetic construct
Qy	181 ATTTGGGAAARGAGAACGGAGGATCGTGGAAAAGTTCACGGTAGCCCTTACGTACATATGGG 240	ORGANISM synthetic construct
Db	241 CGATAGATAGATAACGTGCACGGTCACTTGACGCCACAATTGACGCCACAATTGACGCCA 300	OTHER SEQUENCES: artificial sequences.
Qy	241 CGATAGATAGATAACGTGCACGGTCACTTGACGCCACAATTGACGCCA 300	1 (bases 1 to 46)
Db	301 CTTGACGCTACT 312	AUTHORS Tori, R., Sawant, S.B., Syng, P.K., and Gupta, S.K.
Qy	301 CTTGACGCTACT 312	TITLE Chemically synthesized artificial promoter for realizing high-level expression of introduced gene and method for synthesizing the same.
Db	301 CTTGACGCTACT 312	PATENT: JP 2000139477-A 5 23-MAY-2000;
Qy	301 CTTGACGCTACT 312	JOURNAL COUNCIL OF SCIENTIFIC AND INDUSTRIAL RESEARCH
Db	301 CTTGACGCTACT 312	COMMENT OS Artificial Sequence
		PN JP 2000139477-A/5
		PD 23-MAY-2000
		PF 27-APR-1999 JP 1999119227
		PR 09-NOV-1998 IN 3322/98
		PI RAKESH TORI, SALLY BISHWANATO SAWANTO, PURAJUNNA KUMAR SYNGE,
		PI SHIFU KUMAR, GUPUTA PC C12N15/09,C12N5/10,C12N15/00,C12N5/00
		CC
		KEY Location/Qualifiers
		FT source 1..46 /organism= 'Artificial Sequence'.
		FT Location/Qualifiers
		1..46 /organism= "synthetic construct"
		/mol_type="genomic DNA"
		/db_xref="itaxon:32630"
		ORIGIN
RESULT 2		Query Match 14.7%; Score 46; DB 6; Length 46;
AR427842		Best Local Similarity 10.0%; Pred. No. 0.02;
LOCUS	AR427842 Sequence 1 from patent US 6639065.	Mismatches 0; Indels 0; Gaps 0;
DEFINITION		
ACCESSION	AR427842	
VERSION	AR427842..1 GI:40186826	
KEYWORDS		
SOURCE		
ORGANISM	Unknown. Unclassified.	
REFERENCE	1 (bases 1 to 452)	
AUTHORS	Tuli, R., Sawant, S.V., Singh, P.K. and Gupta, S.K.	
TITLE	Chemically synthesized artificial promoter for high level expression of transgenes	
JOURNAL	Patent: US 6639065-1 28-OCT-2003;	
FEATURES	Location/Qualifiers	
VERSION	1..452	
KEYWORDS	/organism= "unknown" /mol_type= "genomic DNA"	
ORIGIN		
Query Match	99.54; Score 310.4; DB 6; Length 452;	
Best Local Similarity	99.7%; Pred. No. 6.4e-79;	
Matches	311; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
Db	1 GTTCGACCATCATTGAAAGGCCCTCGGTAATACATTGGAAAAAGTGGATAATACCGGA 60	RESULT 4
Qy	1 GTTCGACCATCATTGAAAGGCCCTCGGTAATACATTGGAAAAAGTGGATAATACCGGA 60	E64765 Locus 46 bp DNA linear PAT 31-JAN-2002
Db	1 GTTCGACCATCATTGAAAGGCCCTCGGTAATACATTGGAAAAAGTGGATAATACCGGA 60	DEFINITION Chemically synthesized artificial promoter for realizing high-level expression of introduced gene and method for synthesizing the same.
Qy	61 AAAGAAGATTCACTCATCCAGAAAAGGTGTGGAAAAAGTGTGGATTGGTGGAAAAGTT 120	ACCESSION E64765
Db	61 AAAGAAGATTCACTCATCCAGAAAAGGTGTGGATTGGTGGAAAAGTT 120	VERSION E64765 .1 GI:18628528
Qy	121 CGATCTGACCATCTAGATCGTGGAAAAGTTCACGGTAGCCCTTACGTACATATGGG 180	KEYWORDS synthetic construct
Db	121 CGATCTGACCATCTAGATCGTGGAAAAGTTCACGGTAGCCCTTACGTACATATGGG 180	ORGANISM synthetic construct
Qy	181 ATTTGGGAAAGAGAGGAGGATCGTGGAAAAGTTCACGGTAGCCCTTACGTACATATGGG 240	OTHER SEQUENCES: artificial sequences.
Db	181 ATTTGGGAAAGAGAGGAGGATCGTGGAAAAGTTCACGGTAGCCCTTACGTACATATGGG 240	1 (bases 1 to 46)
Qy	241 CGATAGATAGATAACGTGCACGGTCACTTGACGCCACAATTGACGCCACAATTGACGCCA 300	AUTHORS Tori, R., Sawant, S.B., Syng, P.K., and Gupta, S.K.
Db	241 CGATAGATAGATAACGTGCACGGTCACTTGACGCCACAATTGACGCCACAATTGACGCCA 300	TITLE Chemically synthesized artificial promoter for realizing high-level expression of introduced gene and method for synthesizing the same.
Qy	301 CTTGACGCTACT 312	PATENT: JP 2000139477-A 6 23-MAY-2000;
Db	301 CTTGACGCTACT 312	JOURNAL COUNCIL OF SCIENTIFIC AND INDUSTRIAL RESEARCH
		COMMENT OS Artificial Sequence
		PN JP 2000139477-A/6
		PD 23-MAY-2000
		PF 27-APR-1999 JP 1999119227
		PR 09-NOV-1998 IN 3322/98
		PI RAKESH TORI, SALLY BISHWANATO SAWANTO, PURAJUNNA KUMAR SYNGE,
		PI SHIFU KUMAR, GUPUTA PC C12N15/09,C12N5/10,C12N15/00,C12N5/00

TITLE Direct Submission
JOURNAL Submitted (20-JUN-2002) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

REFERENCE 3 (bases 1 to 114519)

AUTHORS Shauli,S., Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.

TITLE Direct Submission

JOURNAL Submitted (10-OCT-2003) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

REFERENCE 4 (bases 1 to 114519)

AUTHORS Shauli,S., Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.

TITLE Direct Submission

JOURNAL Submitted (05-MAR-2004) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

COMMENT On Mar 5, 2004 this sequence version replaced gi:37202152.

FEATURES

SOURCE 1. 114519 /organism="Medicago truncatula"
 /mol_type="genomic DNA"
 /db_xref="taxon:3880"
 /clone="mth2-27c4"
 /clone_lib="Medicago truncatula BAC library H2"

ORIGIN

Query Match 13.1%; Score 41; DB 8; Length 114519;
 Best Local Similarity 52.7%; Pred. No. 0.95; Indels 0; Gaps 0;
 Matches 89; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Qy 42 AAAAAGTGTGTTAAATACGGAAAAAGAGATTCAATCCAGAAAAGGTGTTGGAAAAGTGT 101
 Db 30972 AAAAAAAAGCTAAAGAAAGAACAGAACAGAGGAGGAGGAGGAGGAGGAGT 30913

Qy 102 GGATTCGCTCGAAAAAAAGTTCGATCTGCACCATCTAGATCCTGGAAAAGTTCACGTAG 161
 Db 30912 GGGTGAATTCAACAAATCAAGTCGAGAGATTGAAACAATTGAAAGAGTTCAACTTGG 30853

Qy 162 CGCTTACGATATGTTGAAAGAGAGCGAGGATCGGT 210
 Db 30852 TGCACAAAGAAAAAACAAATTGAGAAAATGAAATTGTCAAACTT 30804

RESULT 9
AC135605 AC135605 121565 bp DNA linear HTG 13-APR-2004
DEFINITION Medicago truncatula clone mth2-7ml, WORKING DRAFT SEQUENCE.
ACCESSION AC135605
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Eukaryota; Viridiplantae; Streptophytta; Embryophyta; Tracheophyta; Spermatophytina; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilioideae; Trifolieae; Medicago.
REFERENCE 1 (bases 1 to 121565)
AUTHORS Shauli,S., Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.

TITLE JOURNAL
AUTHORS Shauli,S., Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.

TITLE Direct Submission

JOURNAL Submitted (19-OCT-2002) Department Of Chemistry And Biochemistry,

The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

REFERENCE 3 (bases 1 to 121565)
AUTHORS Shauli,S., Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.

TITLE Direct Submission

JOURNAL Submitted (13-APR-2004) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

COMMENT On Apr 13, 2004 this sequence version replaced gi:45861010.

FEATURES

SOURCE 1. 121565 /organism="Medicago truncatula BAC library H2"
 /mol_type="genomic DNA"
 /db_xref="taxon:3880"
 /clone="mth2-7ml"
 /clone_lib="Medicago truncatula BAC library H2"

ORIGIN

Query Match 13.1%; Score 41; DB 2; Length 121565;
 Best Local Similarity 52.7%; Pred. No. 0.95; Indels 0; Gaps 0;
 Matches 89; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Qy 42 AAAAAGTGTGTTAAATACGGAAAAAGAGATTCAATCCAGAAAAGGTGTTGGAAAAGTGT 101
 Db 88922 AAAAAAAAGCTAAAGAAAGAACAGAACAGAGGAGGAGGAGGAGT 88981

Qy 102 GGATTCGCTCGAAAAAAAGTTCGATCTGCACCATCTAGATCCTGGAAAAGTTCACGTAG 161
 Db 88982 GGGTGAATTCAACAAATCAAGTCGAGAGATTGAAACAATTGAAAGAGTTCAACTTGG 89041

Qy 162 CGCTTACGATATGTTGAAAGAGAGCGAGGATCGGT 210
 Db 89042 TGCACAAAGAAAAAACAAATTGAGAAAATGAAATTGTCAAACTT 89090

RESULT 10
AC147540 AC147540 172623 bp DNA linear HTG 06-FEB-2004
DEFINITION Pan troglodytes clone rp43-18c12, WORKING DRAFT SEQUENCE, 32 unorderd pieces.

ACCESSION AC147540
VERSION AC147540_2 GI:42415611
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
MATERIAL BukaYota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE 1 (bases 1 to 172623)
AUTHORS Li,J. and Roe,B.A.

TITLE JOURNAL
AUTHORS Unpublished 2 (bases 1 to 172623)
TITLE Direct Submission
JOURNAL Submitted (09-DEC-2003) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

3 (bases 1 to 172623)
 Li, J. and Roe, B.A.
 Direct Submission
 Submitted (06-FEB-2004) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 On Feb 6, 2004 this sequence version replaced gi:39573794.
 COMMENT: Genome Center
 Center: Department Of Chemistry And Biochemistry
 The University Of Oklahoma
 Center code:DONOR

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 32 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

*	1	2143:	contig of 2143 bp in length
*	2144	2243:	gap of unknown length
*	2244	4771:	contig of 2528 bp in length
*	4772	4871:	gap of unknown length
*	4872	7178:	contig of 2307 bp in length
*	7179	7278:	gap of unknown length
*	7279	9321:	contig of 2043 bp in length
*	9322	9441:	gap of unknown length
*	9442	12171:	contig of 2750 bp in length
*	12172	12271:	gap of unknown length
*	12272	14443:	contig of 2172 bp in length
*	14444	14543:	gap of unknown length
*	14544	14736:	contig of 3193 bp in length
*	14737	17836:	gap of unknown length
*	17837	21330:	contig of 3514 bp in length
*	21351	21450:	gap of unknown length
*	21451	24909:	contig of 3459 bp in length
*	24910	25009:	gap of unknown length
*	25010	28464:	contig of 3155 bp in length
*	28465	28544:	gap of unknown length
*	28565	32065:	contig of 3501 bp in length
*	32066	32165:	gap of unknown length
*	32166	36136:	contig of 4031 bp in length
*	36197	36296:	gap of unknown length
*	36297	41563:	contig of 5267 bp in length
*	41564	41663:	gap of unknown length
*	41664	45475:	contig of 3812 bp in length
*	45476	45575:	gap of unknown length
*	45576	48659:	contig of 3114 bp in length
*	48650	48789:	gap of unknown length
*	48790	53126:	contig of 4937 bp in length
*	48791	53827:	gap of unknown length
*	53727	60347:	contig of 6621 bp in length
*	60348	60547:	gap of unknown length
*	60548	67085:	contig of 6538 bp in length
*	67086	67185:	gap of unknown length
*	67186	71328:	contig of 4443 bp in length
*	71329	71728:	gap of unknown length
*	71729	76114:	contig of 4586 bp in length
*	76313	76414:	gap of unknown length
*	76415	82210:	contig of 6396 bp in length
*	82211	82910:	gap of unknown length
*	82911	87615:	contig of 4705 bp in length
*	87616	87715:	gap of unknown length
*	87716	94195:	contig of 6580 bp in length
*	94296	94495:	gap of unknown length
*	94396	98210:	contig of 4630 bp in length
*	99025:	99025:	gap of unknown length
*	99026	99125:	contig of 5676 bp in length
*	99126	104801:	gap of unknown length
*	104802	104901:	gap of unknown length
*	104903	111798:	contig of 6897 bp in length
*	111799	112898:	gap of unknown length
*	112899	113365:	contig of 9467 bp in length

FEATURES Location/Qualifiers
 1. .172623
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="Laxom:9598"
 /clone="rp43-18c12"
 /clone_lib="RPCI - 43 Male Chimpanzee BAC Library"

 ORIGIN
 source
 1. .172623
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="Laxom:9598"
 /clone="rp43-18c12"
 /clone_lib="RPCI - 43 Male Chimpanzee BAC Library"

 Query Match 13.0%; Score 40.6; DB 2; Length 172623;
 Best Local Similarity 48.1%; Pred. No. 1.3;
 Matches 115; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

 Qy 40 GGAAAAAGTGGTAATAGGAAAAGAGAATCATTCCAGAAAAGCTGGAAAGTT 99
 Db 107763 GGAGACTTCAGATCAAAAGGTGGAAATCTGGCCATCGGCATGGAGCT 107704

 Qy 100 GTGGATTGGTGGAAAAGTGGCATCTGACCATCTTAGATGTTGAAAAACTTCAGTT 159
 Db 107703 CAGCCTAGCATGCCAACCTCTGCAAAACAGTCTTGCA 107644

 Qy 160 AGGGCTTAGTACATAGTGTGATTGTGAAAAGAACGAGGGCTGGCTAAAGAGA 219
 Db 107643 CATTTTAAGTGAATTGAGATTGAGGTGTGAAAGCAGATCATTTCTAGCTCAAATG 107584

 Qy 220 AGCTTGACGTGACGGTGA-CATAGATACTAGATGACGGCTTGACGCCA 278
 Db 107583 AATTAAAGCAAGATACTGGAAATGTAGTTGATCCAGGTTCTACGGTGTATTGACCCA 107525

 RESULT 11
 AC145064 AC145064 197004 bp DNA linear HTG 23-APR-2004
 DEFINITION Pan troglodytes clone rp43-45h9, WORKING DRAFT SEQUENCE, 11
 unorderd pieces.

 ACCESSION AC145064
 VERSION 6 GI:46518614
 KEYWORDS HTG; HTGS PHASE1; HTGS-DRAFT
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes
 Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Buteleostomi;
 1 (bases 1 to 197004)
 REFERENCE Lau,J.C.Y. and Roe,B.A.
 AUTHORS Pan troglodytes BAC Clone rp43-45h9
 TITLE Unpublished
 JOURNAL 2 (bases 1 to 197004)
 REFERENCE Lau,C.C.Y. and Roe,B.A.
 AUTHORS Direct Submission
 TITLE Submitted (05-JUN-2003) Department Of Chemistry And Biochemistry,
 JOURNAL The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019 USA
 REFERENCE 3 (bases 1 to 197004)
 AUTHORS Lau,C.C.Y. and Roe,B.A.
 TITLE Direct Submission
 JOURNAL Submitted (23-APR-2004) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019 USA
 COMMENT On Apr 23, 2004 this sequence version replaced gi:40018761.
 Center: Department Of Chemistry And Biochemistry
 Center code:UOKNOR
 Center: Department Of Chemistry And Biochemistry
 Center code:UOKNOR

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 11 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

*	*	1	3053:	contig of 3053 bp in length
*	*	3054	3153:	gap of unknown length
*	*	3154	6132:	contig of 3029 bp in length
*	*	6183	6283:	gap of unknown length
*	*	6283	1030:	contig of 4048 bp in length
*	*	1030:	10430:	gap of unknown length
*	*	10311	10431:	gap of unknown length
*	*	10431	17594:	contig of 7162 bp in length
*	*	17593	17632:	gap of unknown length
*	*	17693	27462:	contig of 9770 bp in length
*	*	27463	27563:	gap of unknown length
*	*	27563	41677:	contig of 14115 bp in length
*	*	41578	41777:	gap of unknown length
*	*	41778	57017:	contig of 15240 bp in length
*	*	57018	57117:	gap of unknown length
*	*	57118	76985:	contig of 19868 bp in length
*	*	76986	77085:	gap of unknown length
*	*	77085	102632:	contig of 25597 bp in length
*	*	102633	102732:	gap of unknown length
*	*	102783	135930:	contig of 33128 bp in length
*	*	135941	136010:	gap of unknown length
*	*	136011	197004:	contig of 60994 bp in length.

LOCATIONS

NIGHT

Query	Match	Score	40.6;	DB	2;	Length	197004;
Best Local Similarity	13.0%						
Pred. No.	48.1%						
Matches	115;	Conservative	0;	Mismatches	124;	Indels	0;
						Gaps	0;
40	GGAAAAAGTGGTAATACGGAAAAGAAGATTCACTCCAGAAAAAGTGGAAAAAGTT	99					
140399	GGAGACTTCAGATAAAAGGGTAGGGAATCTTGCCATCTGGRATGGACGCT						140458
100	GTGGATTGGTGGAAAAGTTCGATCTGACCATTCTTGAATCCTCGAAAGTTCACGTT	159					
140459	CAGCCTAGATGGCCAAGCTGGAACTCTTCGACAAACAGGTTGTAGCTTCTCAT						140518
160	AGGCCATTAGTACATATGNGATGTTGGAAAAGAAGGGGCACTGGTGGAAAAAGACA	219					
140519	CATTITIVAGTGAATTGAGGACTGTTGGAAACAGCCAGATCATGTTTCAGCTCAAATG						140578
220	AGCTTGAGCTGTAGCTGAGCATAGATACCTGACCGTCACTTGACGCGCA	278					
140529	TATTAAAGCTTAAAGATAACGCAAAATGTTGATGTTGATGTTGATGTTGATGTTG						140637

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	SOURCE	GENE	MINICODE
58	18	78	7					
58	18	78	7					

J02739
/gene="P0533B06_1"
/note="probably inactive due to 5' exon missing in CDS
probably inactive due to including stop codon(s) in CDS
pseudogene, SERK2 protein"
/pseudo

S237 . 6752
/gene="P0533B06_2"
S237 . 6752
/gene="P0533B06_2"
/note="probably inactive due to 5' exon missing in CDS
pseudogene, SERK1 protein"
/pseudo
complement(14120..17601)
/gene="P0533B06_3"

gene
misc_feature

RESULT 12
AP004619.c
ORYZUSATIVUS
Orza sativa (iaponica cultivar-group)
134442 bp
DNA
genomic DNA
chromosome 8.
linear
PLN 03-DEC-2003

complement (14120. . 17601)
 /gene="P0583B06_3"
 /note="probably inactive due to including frameshift(s) in CDS
 pseudogene, RNA-directed DNA polymerase homolog
 /pseudo
 complement (18030. . 19769)
 /gene="P0583B06_4"
 complement (join(<18030. . 18299,18388. . 19352,19433. . 19546
 19646. . >19769))
 /note="start and end point are not identified"
 complement (join(18030. . 18299,18388. . 19352,19433. . 19546,
 19646. . 19769))
 /gene="P0583B06_4"
 /note="predicted by GeneMark,hmm etc."
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="BAD03229_1"
 /db_xref="GI:38636869"
 /translation="MARPQDLANLFDISRCIPSTAVREBEGILRSAAVVRITGVY
 LIGCWNINSEPPDIAFKHQLQLEELNQNSPDPEFEDLGSSRGLDILKRLTIVQ
 LTHDQVLISSPPEPMWPGYSPESQSGMIISSAVTEVPKRHDIIWISVGKQDCQRA
 PRHVAARLGPGRNRSHSSRSRPAEHRRPSGNDVPPARTAATPRRSSSRPDRHGG
 SGGYQRATSPSPRERTSPRNTSPPPAPRPLQISMPPELSEBTFRPESQ
 TYPVPEFTATVQVDQDANSFAGDDMMAPLSVDPMLQSLPPEGLGNAAPEPQS
 OPOAQPPGELARCKRARIORSARIAKKVVSRLAPAKPGRAQEDLMKKAGLISPDKKK
 TDQAMTRKVLKFLPQLSARVIDAATVLAQGAIQGDKKPF"
 complement (19819. . 21046)
 /gene="P0583B06_5"
 complement (join(<19819. . 20120,20296. . 20499,20584. . 20686
 21038. . >21046))
 /gene="P0583B06_5"
 /note="start and end point are not identified"
 complement (join(19819. . 20120,20296. . 20499,20584. . 20686,
 21038. . 21046))
 /gene="P0583B06_5"
 /note="predicted by GeneMark,hmm etc."
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="BAD03230_1"
 /db_xref="GI:38636870"
 /translation="MNLATSKFHDPEFTPPALPTGHRHCDPSTHTAGJHSGESEED
 ISPPPPDGGSARASVSGCSWRWLLRLRDGMWLPAGSAASSSSRYSSYSASC
 LCAAYLRKLSTLSSLTTYQOASRFTNTSLSIQLPKAGSSWSNSLKGROLEDVFAASPR
 LTALTRVETHPGASSALPPLATVLAQGILQFRNRNRFTRSSPCLLIA"
 21647. . 29582
 /gene="P0583B06_6"
 complement (join(<21647. . 22527,27443. . >29582))
 /gene="P0583B06_6"
 /note="start and end point are not identified"
 join(21647. . 22527,27443. . 29582))
 /codon_start=1
 /product="putative NBS-LRR disease resistance protein
 homologue"
 /protein_id="BAD03231_1"
 /db_xref="GI:38636871"
 /translation="MLSYDVKREFOEIDRSAMVGALASASTGVMNSVIAKLSKLDEEV
 AKLGKVQQQIAPIRDELBMNATLRLVALDEEEDLDPVWKRDKVRELTEDDCIDS
 FEVTRISQERGSGLIKIIIRKLKQKLRHAFANQIEVVEESRKRQYDILK
 PMSVSSAATPTIDRPLAYPEVQDOKLVKGMPRENTIENTKSDRSRDRSLRKHIVSVG
 CGGLGKTTIANQVKPKEHQDSDAFTVSURNPDIKKLRLMKEVNSLDNTQPWSNP
 SKCSYHGXYHNEPLSPFDSKRLPQFQGSENLYYPLPEEVNGVSKMKGCGGILPVA
 ITISSLADQADENTVTAIGSALAKDPMAGNMRLJFSYVDPVHLTRCTLILV
 VACRHVDLILDFETCKAABENFTLDTDPEFRVYFONNENKVYMATNL
 SKVNRSLTIFGYFKVTPSLLDFKVLVRLDQKLOQHHLTCIEMLHLKTLGSL
 RYTTELPGKLGRLLETIETMWSLPSATFTQFVRLLNHDTTFSQPCMHTGJMSL
 SPCFQDCEPTEVYSSMPLTQKLRKJRVTFQFVRLLNHDTTFSQPCMHTGJMSL

SOURCE	Oryza sativa (japonica cultivar-group) Oryza sativa (japonica cultivar-group) Eukaryota; Viridiplantae; Streptophytina; Embryophytina; Tracheophytina; Spermatophytina; Magnoliophytina; Liliopsida; Poales; Poaceae;
REFERENCE	
AUTHORS	Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE	Orzya sativa nipponbare (GA3) genomic DNA, chromosome 8, BAC clone OJ1120 C08 Published Only in Database (2001)
JOURNAL	
REFERENCE	Sasaki,T., Matsumoto,T. and Yamamoto,K.
AUTHORS	
TITLE	Direct Submission
JOURNAL	Submitted (09-JUL-2001) Takaji Sasaki, National Institute of Biological Sciences, Rice Genome Research Program; Kamionda 2-1-2, Tsuchiura, Ibaraki 305-8602, Japan (E-mail: t-sasaki@nias.affrc.go.jp URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT	On Dec 2, 2003 this sequence version replaced gi:33235570. Genes were predicted from the integrated results of the following: GENSCAN (http://CCR-001.mit.edu/GENSCAN.html), FGENESH (http://www.softberry.com/), GeneMark (http://open.biology.gatech.edu/GeneMark/), GlimmerM (http://www.tigr.org/cdb/glimmer/Glmr/), SplicePredictor (http://rgp.dna.affrc.go.jp/RicefMM/), SplicePredictor (http://bioinformatics.iastate.edu/cgi-bin/sp_cgi), SLM4 (http://globin.cse.psu.edu/html/docs/sim4.html), Gap2 (http://www.tigr.org/software/glimmerm/), BLASTN and BLASTX. The genomic sequence was searched against NCBI NonRedundant Protein database, nr (Ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP or DBJU. Protein homologies of the coding regions were searched against NCBI NonRedundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no.
FEATURES	<p>A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence.</p> <p>The orientation of the sequence is from -21M13 to M13rev of the BAC clone. This sequence of OJ1120_C08 clone has an overlap with P0577G06 (DDBJ: AP004619) clone at 5', end and with P0577G06 (DDBJ: AP004590) clone at 3', end. The sequence was generated by combining Monsanto and RGP-Japan sequencing data. Detailed information on overlap and assembly sequencing data, together with annotation of this entry is available at http://rgp.dna.affrc.go.jp/genomeseq.html.</p>
source	<p>1. .150300 Location/Qualifiers (organism="Oryza sativa (japonica cultivar-group)" mol_type="genomic DNA" cultivar="Nipponbare" db_xref="Caxon:39947" chromosome="8" clone="OJ1120_C08" S270..8150 Gene="OJ1120_C08_1" S270..8150_Pseudo Gene="OJ1120_C08_1" note="probably inactive due to 3' exon missing in CDS pseudogene, MLA6 protein" 10283..12816_Pseudo Gene="OJ1120_C08_2" 10283..12816</p>
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
FEATURES	
source	

gene	AYHDLRPAAKREAAARRDVFVFFVNQGWRGTTGARERFLNTAGFPESEVIRVLADPOLEPMT ROADVQGLQKQKARHEDLSQGMKQITLDSLAEAVVTSCTCDASPPPPPEE EPSSAAQDGCGMAQNGSAMRLSNIVAFDHVRURRSPPPTSHLHVRAELG" complement	join(29680 . 29827 , 31003 . 31098 , 32215 . 32219 , 33217 . 33312) . 9			
misc_feature	/gene="OJ1120_C08_8"				
gene	/note="hypothetical ORF predicted by GENSCAN this category is not included in IRGSP standard" join(37583 . 37663 . 37864 . 38022)				
mRNA	/gene="OJ1120_C08_10"				
gene	/note="hypothetical ORF predicted by GENSAN this category is not included in IRGSP standard" complement				
mRNA	/gene="OJ1120_C08_10"				
CDS	/note="start and end point are not identified" complement				
gene	/note="predicted by FGENESH etc." (codon_start=1 /product="hypothetical protein" /protein_id="BAD03040_1" /db_xref="GI:386326799" /translation="MYYEAAAVASSLKLKGQQQORGHKGSAARSHEHQVSIEKQQQQIG CRILWAPRRTYTAGSKR" complement				
mRNA	/gene="OJ1120_C08_11"				
gene	/note="hypothetical ORF predicted by GlimmerM this category is not included in IRGSP standard" complement				
gene	/note="supported by full-length cDNA (S): AK105561" complement				
CDS	/note="full-length cDNA (S): AK105561" Query Match 12.6% ; Score 39.2 ; DB 8 ; Length 150300 ; Best Local Similarity 50.0% ; Pred. No. 3.2 ; Matches 98 ; Conservative 0 ; Mismatches 98 ; Indels 0 ; Gaps 0 ;				
Qy	29 AATACCATGTGGAAAAAGTGTGTAATACTGGAAAAGAGATTCACTATCCAGAAAAGCT 88				
Db	83780 AAAAAGAAAGTAATAAGATTGTCATCTGAAGATTGAGATAGCTAACAAAAGG 83721				
Qy	89 GTGCAAAGTGTGCAAAAGTGTGCAATGTCGACCATCTCATGATGTGGAA 148				
Db	83720 GAAACAAAGTGTGCAACAAATGTTGACTTGTGAATAA 83561				
Qy	149 AAGTTCAGTTAGGCCTACATATGGATGTGGATGTTGAAAGAAGAACGGGGCATCG 208				
Db	83660 AAGGAAATAAACAGGGCATTGTTGCAATGAAATAAATGGAG 83601				
Qy	209 GTGAAAGAAGCTT 224				
Db	83600 GATGGAAATAAGCTT 83558				
RESULT 14					
AC112060/c					
LOCUS	AC112060	182932 bp	DNA	linear	HTG 15-NOV-2002

DEFINITION	Rattus norvegicus clone CH230-180B21, *** SEQUENCING IN PROGRESS
VERSION	***, 3 unordered pieces.
ACCESSION	AC112060
KEYWORDS	HTG; HTGS; DRAFT; HTGS_ENRICHED.
ORGANISM	Rattus norvegicus
ORGANISM	Rattus norvegicus
ORGANISM	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
REFERENCE	Rattus
AUTHORS	1 (bases 1 to 182932)
Muzny, D. Marie., Metzker, M. Lee., Abramzon, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baden, H., Baldwin, D., Baudaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesari, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Andrade, C., Dederich, D., Delgado, O., Denison, S., Dernamo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, B., Eaves, K., Egan, A., Escott, M., Eugene, C., Evans, C.A., Failes, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, P., Garcia, A., Garner, T., Gebregheorgis, B., Gill, R., Grady, M., Guerra, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlik, P., Hayes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgeson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., Khan, Z., King, L., Kovari, C., Kowis, C., Kraft, C.L., Lebow, H., Levant, J., Lewis, L., Li, Z., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Loresniewa, L., Louisedeg, H., Lozano, R.J., Lu, X., Ma, J., Mahmud, M., Mallard, M., Mapu, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenan, E., Milesavajevic, A., Miner, G., Minja, E., Montenayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakelueh, O., Okwonou, G., Olarnpusang, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, J.L., Pfannkoch, C., Plopper, F., Ponderex, A., Popovic, D., Primus, E., Pu, L.-L., Puazzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, T., Rojas, A., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sission, I., Sitter, C.D., Smajis, D., Sneed, A., Sodergren, B., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingye, A., Trejos, Z., Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Wilson, R., Wleczcy, R., Woodin, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederauer, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Weinstock, G.A.	
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 182932)
AUTHORS	Worley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (19-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3 (bases 1 to 182932)
AUTHORS	Rat Genome Sequencing Consortium.
TITLE	Direct Submission
JOURNAL	Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT	Baylor Plaza, Houston, TX 77030, USA On Nov 15, 2002 this sequence version replaced gi 22856848; the sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly ('contig=contig'). Within each contig=contig, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig=contig that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.	Qy	197	ACGGAGGCATCGTGGAAAAAGAA	220
		Db	179898	ATGGAACCAACCTGTGATAGAATGAA	179875
RESULT 15		AC108612	185618	DNA linear	HTG 20-NOV-2002
	LOCUS	Rattus norvegicus clone CH230-293B15, WORKING DRAFT SEQUENCE, 2			
DEFINITION	unordered pieces.				
ACCESSION	AC108612				
VERSION	AC108612.5				
KEYWORDS	GT;25129237				
SOURCE	HTGS PHASE1; HTGS_FULLTOP.				
ORGANISM	Rattus norvegicus (Norway rat)				
MAMMALIA;	Eutheria;	Rodentia;	Sciurognathi;	Muridae;	Murinae;
RATTOUS;	Metazoa;	Chordata;	Craniata;	Vertebrata;	Euteleostomi;
RATTUS;					
1 (bases 1 to 185618)					
REFERENCE					
AUTHORS	Muzny, D. Marie, Metzker, M. Lee, Abramzon, S., Adams, C., Alder, J., Allen, C., Allen, H., Alstroem, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, J., Baca, E., Baden, H., Baldwin, D., Bandara-naika, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chave, D., Chen, G., Chen, Y., Chen, Z., Cleveland, C., Cockrell, R., Cox, C., Coyne, M., Cree, A., D'Souza, L., Davila, M.L., Davis C., Davy-Carroll, L., De Andrade, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Baves, K., Egan, A., Escott, M., Eugene, C., Evans, C.A., Fallis, T., Fan, G., Hernandez, R., Hinze, S., Hladun, S. L., Hodges, A., Hooges, M., Hollings, B., Howell, S., Huldyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, R., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowalski, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Loresuhawa, L., Louiselard, R.J., Lozad, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindarne, M., Mahmoud, M., Mallory, K., Mangum, A., Mangum, B., Mapua, P., McNeill, T.Z., Meinen, E., Mawhinney, S., McLeod, M.P., McNeill, K., Miner, G., Montemayor, J., Moore, S., Milosavljevic, A., Miner, G., Minja, B., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidas, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoleleme, O., Okonou, G., Olarnpunsgagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, J., Pfannkoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, B., Pu, L.-L., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, T., Rojas, A., Rose, M., Ross, R., Ruiz, S.J., Sanders, W., Savery, G., Scheerer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Snee, A., Sodergren, E., Song, X.-Z., Sosa, J., Steimle, M., Strong, R., Sutcliffe, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, R., Wei, X., White, F., Williams, G., Wilson, R., Wlezy, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yaku, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederauer, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.				
FEATURES					
source	/organism="Rattus norvegicus" /mol_type="genomic DNA" /db_xref="taxon:10116" /clone="CH230-180BB21"				
misc_feature	1..1190				
misc_feature	177608 ..177536				
ORIGIN	/note="wgs_contig" /note="wgs_contig"				
Qy	17	AGGGCCCTCGGTAATACATTTGTTGAAAGTTGTGAAATACGGAAAAAGAGATTCTCA	76	Query Match Score 12.6%; Length 182932;	
		180078	AAAGACATGAGAACACTCCTTAGGAAACAGAAAACATTATAACAGTAACTAGTAGAACCC	180019	Best Local Similarity 49.5%; Pred. No. 3.2;
					Matches 101; Conservative 0; Mismatches 103; Indels 0; Gaps 0;
Db	77	TCCAGAAAGTGTGAAAGTGTGTGCTGCTGAAAGTCGATCTCT	136		
Qy	180018	TACAGAGGAATCCCAAATCCGTGAGAACATGCCATGAAATGCGTGAAGAACATCCA	179959		
Qy	137	AGATCCTGGAAAAAGTTCAGTTGCGCTTACGTACATATGAGATGAGAG	196	JOURNAL	
Db	179958	GGAAACACAAACAGTGAAGGAACCTAAATGAAATGAAAGAACAC	179899	AUTHORS	
					TITLE
					Direct Submission
					Unpublished
					2 (bases 1 to 185618)
					Worley, K.C.
					Direct Submission

JOURNAL Submitted (31-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 REFERENCE 3 (bases 1 to 185618)
 AUTHORS Rat Genome Sequencing Consortium.
 TITLE Direct Submission
 JOURNAL Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 COMMENT On Nov 20, 2002 this sequence version replaced gi:2285635. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

 Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GPOO

Center clone name: CH230-293B15

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 178343 bases at least Q40

Consensus quality: 173378 bases at least Q30

Consensus quality: 180026 bases at least Q20

Estimated insert size: 180394; sum-of-contigs estimation

Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 2 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 184363: contig of 184363 bp in length

* 184463: gap of unknown length

* 184464 185618: contig of 1155 bp in length.

FEATURES

source

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-293B15"

misc_feature 1..1155

/note="wgs_contig"

misc_feature 183070..183363

/note="wgs_contig"

ORIGIN

Query Match 12.6%; Score 39.2; DB 2; Length 185618;
 Best Local Similarity 49.5%; Pred. No. 3.3;
 Matches 101; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

Qy 17 AAGGGCCCTCGGTAAATACATTGTTGAAAGATGTCGTTACGGAAAAGAAGATTCA 76

Db 1383350 AAAGGATGAGAACTCCCTAGGAAACAGAAAACATTAAACAAGTGAAGGCC 138409

Qy 77 TCCAGAAAAGTGTGAAAGTGTGTGAAAGTGTGTGAAAGTGTGTGAAATGAGATCTGACATCCT 136

Large blank (USP)

Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 00:53:27 ; Search time 474.353 Seconds
(without alignments)

3893.639 Million cell updates/sec

Title: US-10-814-858A-1

Perfect score: 312

Sequence: 1 gtgcaccatcatgtaaagg.....tgacgccacttgaacctact 312

Scoring table: IDNTNT NUC Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0% Maximum Match 100%

Listing first 45 summaries

Database :	N_Geneseq_16Dec04 :*
1:	Geneseqm1980s:*
2:	Geneseqm1990s:*
3:	Geneseqm2000s:*
4:	Geneseqm2001s:*
5:	Geneseqm20010s:*
6:	Geneseqm2002as:*
7:	Geneseqm20020s:*
8:	Geneseqm2003as:*
9:	Geneseqm2003bs:*
10:	Geneseqm2003cs:*
11:	Geneseqm2003ds:*
12:	Geneseqm2004as:*
13:	Geneseqm2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	
1	310.4	99.5	452	3 AAA28449	Aaa28449 Artificial	
2	46	14.7	46	3 AAA28438	Aaa28438 Synthetic	
3	46	14.7	46	3 AAA28437	Aaa28437 Synthetic	
4	36.8	11.8	110000	1 ACN5090_1	Continuation (2 of Abi02730 Drosophili	
5	34.4	11.0	7450	4 ABLO730	Abi02730 Drosophili	
C	6	33	10.6	4074	10 ADBs9083	C. neotor
7	32.6	10.4	468	6 ABK76259	Abk76259 Bacillus	
8	32.6	10.4	110000	6 ABQ69245_01	Continuation (2 of Abi067197 Listeria	
9	32.6	10.4	110000	6 ABQ67197_00	Abi05935 S. pneumo	
10	32.4	10.4	495	10 ABX05935	Abx05935 S. pneumo	
C	11	32.4	10.4	1096	2 AAT98670	Aat98670 DNA encod
C	12	32.4	10.4	1727	2 AAZ95371	Aaz95371 S. Pneumo
C	13	32.4	10.4	2992	4 AA166078	Aa166078 Spinacia
C	14	32.4	10.4	6380	6 ABN80109	Abn80109 Human che
C	15	32.4	10.4	11303	2 AAV52248	Aav52248 Streptoco
C	16	32.4	10.4	65454	12 ADN01773	Adn01773 Human hun
C	17	32.4	10.4	83517	13 ABD12581	Abd12581 Human can
C	18	32.4	10.4	110000	10 ABS56454_01	Continuation (2 of Abi032561 Human can
C	19	32	10.3	110000	6 ABA03401_22	Continuation (23 o
C	20	31.8	10.2	722	6 ABQ68303	Abq68303 Listeria

ABQ70825 Listeria
Acc28221 Prokaryot
Aaa12022 Human sec
Abx56693 Arabidops
Abi109683 Drosophil
Abi119785 Drosophil
Aar09296 Nucleotid
Aad05761 Arabidops
Aad05787 Arabidops
Aad06648 A. thalia
Abk65208 Arabidops
Add55841 Thalecress
Add30760 Plant yie
Aci41682 Plant tra
Ado2050 Thalecress
Ado03392 Thalecress
Abi109682 Drosophil
Abi119784 Drosophil
Abk40014 Human che
Aab20588 Human met
Abd33020 Mouse can
Aah00080 Enterococ
Ada71938 Rice gene
Abi34132 Human imm
Aba21231 Human ner

ALIGNMENTS

RESULT 1

ID	AA28449 standard; DNA; 452 BP.
XX	AC
XX	DT 29-AUG-2000 (first entry)
XX	DE Artificial synthetic promoter for high level plant gene expression.
XX	FH Artificial promoter; strong; weak; transgene expression; plant; ss.
XX	OS Synthetic.
XX	Key Location/Qualifiers
FT misc_signal	1..120 /*tag=^a
FT misc_signal	/note= "domain II"
FT misc_signal	121..142 /*tag=^b
FT misc_signal	/note= "domain III"
FT misc_signal	143..220 /*tag=^c
FT misc_signal	/note= "Domain III"
FT misc_signal	221..266 /*tag=^d
FT misc_signal	/note= "Domain I"
FT misc_signal	267..307 /*tag=^e
FT misc_signal	/note= "Minimal domain Ib"
FT misc_signal	308..324 /*tag=^f
FT misc_signal	/note= "Minimal domain Ia"
FT misc_signal	317..324 /*tag=^g
FT misc_signal	/note= "plant consensus TATA signal"
5' UTR	351..452 /*tag=^h

PF 25-FEB-1999; 99EP-00301419.
 XX
 PR 09-NOV-1998; 98IN-DE003322.
 XX
 PA (COUL) CSIR COUNCIL SCI IND RES.
 XX
 PI Tuli R, Sawant SV, Singh PK, Gupta SK;
 XX
 DR; 2000-341712/30.
 XX
 PT New chemically synthesized artificial promoter, useful high level
 expression of transgenes in different organisms.
 XX
 PS Claim 13; Page 11; 40pp; English.
 XX
 CC Chemically synthesized artificial promoters are new and comprise a DNA
 sequence designed for a targeted level and pattern of gene expression by
 strategically putting together several signature sequences identified by
 sequence alignment and statistical analysis of a large database
 constructed for this purpose. A method for chemically synthesizing an
 artificial promoter for expressing genes at a desired level in different
 organisms is also claimed. The high level expression in a plant using
 such an artificial promoter (e.g. AAA28449) can be measured comprising
 polyethylene Glycol (PEG) mediated transformation of plant protoplasts as
 well as biolistic mediated transformation of plant tissues including
 root, stem, intact leaf tissue followed by transient GUS assay to compare
 with a natural CaMV 35S promoter showing the desired level of activity.
 The promoter is useful for high level expression of transgenes in
 different organisms and for testing high level gene expression in plants
 (claimed). The promoter is biologically active and is efficient and can
 be synthesized to express in even the most complex organisms.
 XX
 SQ Sequence 452 BP; 157 A; 88 C; 95 G; 112 T; 0 U; 0 Other;
 XX
 Query Match 99.5%; Score 310.4; DB 3; Length 452;
 Best Local Similarity 99.7%; Pred. No. 2e-87;
 Matches 311; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 XX
 Qy 1 GTGACCCATCATTCATGGCTGGTAAATACCATTGCGAAAAGTGTTGATAACGA 60
 Db 1 GTGCACCATCATTCATGGAAAGGCCCTGGTAAATACCAATTGGAAAAAAGTGGTAATACGA 60
 XX
 CC A chemically synthesized promoter can comprise a conserved domain I as
 shown here for high level expression of genes. Chemically synthesized
 artificial promoters are new and comprise a DNA sequence designed for a
 targeted level and pattern of gene expression by strategically putting
 together several signature sequences identified by sequence alignment and
 statistical analysis of a large database constructed for this purpose. A
 method for chemically synthesizing an artificial promoter for expressing
 genes at a desired level in different organisms is also claimed. The high
 level expression in a plant using such an artificial promoter (e.g.
 AAA28449) can be measured comprising polyethylene Glycol (PEG) mediated
 transformation of plant protoplasts as well as biolistic mediated
 transformation of plant tissues including root, stem, intact leaf tissue
 followed by transient GUS assay to compare with a natural CaMV
 35S promoter showing the desired level of activity. The promoter is useful
 for high level expression of transgenes in different organisms and for
 testing high level gene expression in plants (claimed). The promoter is
 biologically active and is efficient and can be synthesized to express in
 even the most complex organisms
 XX
 SQ Sequence 46 BP; 11 A; 11 C; 13 G; 11 T; 0 U; 0 Other;
 XX
 Query Match 14.7%; Score 46; DB 3; Length 46;
 Best Local Similarity 10.0%; Pred. No. 0.00017;
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 Qy 221 GCTTGTACGGCTTACGGCTTACGGTACGGTACGGTACGGT 266
 Db 1 GCTTGTACGGCTTACGGCTTACGGTACGGTACGGTACGGT 46
 XX
 RESULT 3
 AAA28437

RESULT 2
 AAA28438 Standard; DNA; 46 BP.
 XX
 AC
 XX
 DT 29-AUG-2000 (first entry)
 XX
 DE Synthetic promoter conserved domain I.
 XX

ID AAA28437 standard; DNA; 46 BP.
 XX
 AC AAA28437;
 XX
 DT 29-AUG-2000 (First entry)
 XX
 Synthetic promoter minimal domain (b) for high level gene expression.
 DE
 XX
 KW Artificial promoter; strong; weak; transgene expression; plant; ss.
 XX
 OS Synthetic.
 XX
 KEY CAAT_signal
 FT 1..5
 FT /*tag= a
 FT /note= "conserved"
 FT misc_signal
 FT 6..10
 FT /*tag= b
 FT /note= "conserved TGACG box"
 FT CAAT_signal
 FT 11..16
 FT /*tag= c
 FT /note= "conserved"
 FT misc_signal
 FT 17..21
 FT /*tag= d
 FT /note= "conserved TGACG box"
 FT CAAT_signal
 FT 22..27
 FT /*tag= e
 FT /note= "conserved"
 FT misc_signal
 FT 27..31
 FT /*tag= f
 FT /note= "conserved TGACG box"
 FT CAAT_signal
 FT 32..36
 FT /*tag= g
 FT /note= "conserved"
 FT misc_signal
 FT 37..41
 FT /*tag= h
 FT /note= "conserved TGACG box"
 FT CAAT_signal
 FT 42..46
 FT /*tag= i
 FT /note= "conserved"
 XX EP1002869-A1.
 XX
 PD 24-MAY-2000.
 XX
 PF 25-FEB-1999; 99EP-00301419.
 XX
 PR 09-NOV-1998; 98IN-DE003322.
 XX
 PA (COUNCIL CSIR COUNCIL SCI IND RES.
 XX
 PI Tuli R., Sawant SV, Singh PK, Gupta SK;
 XX
 DR WPI; 2000-341712/30.
 XX
 PT New chemically synthesized artificial promoter, useful high level
 PT expression of transgenes in different organisms.
 XX
 PS Claim 4; Page 15; 40pp; English.
 XX
 CC A chemically synthesized promoter can comprise a minimal domain (b) as
 CC shown here for high level expression of genes. This domain has a number
 CC of conserved sequences interspersed with a TGACG box. Chemically
 CC synthesized artificial promoters are new and comprise a DNA sequence
 CC designed for a targeted level and pattern of gene expression by
 CC strategically putting together several signature sequences identified by
 CC sequence alignment and statistical analysis of a large database
 CC constructed for this purpose. A method for chemically synthesizing an
 CC artificial promoter for expressing genes at a desired level in different
 CC organisms is also claimed. The high level expression in a plant using
 CC such an artificial promoter (e.g. AAA8449) can be measured comprising
 CC polyethylene glycol (PEG) mediated transformation of plant protoplasts as
 CC well as biolistic mediated transformation of plant tissues including

CC root, stem, intact leaf tissue followed by transient GUS assay to compare
 CC with a natural CamV 35S promoter showing the desired level of activity.
 CC The promoter is useful for high level expression of transgenes in plants
 CC different organisms and for testing high level gene expression in plants
 CC (claimed). The promoter is biologically active and is efficient and can
 CC be synthesized to express in even the most complex organisms
 XX SQ Sequence 46 BP; 13 A; 16 C; 8 G; 9 T; 0 U; 0 Other;
 XX
 Query Match 14.7%; Score 46; DB 3; Length 46;
 WP Best Local Similarity 100.0%; Pred. No. 0.00017;
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 267 CCACCTGAGCCATTGAGCCTTGAAGCTACT 312
 Db 1 CCACCTGAGCCACCATTGAGCCTAATGAGCTACT 46
 RESULT 4
 ACN45090_1
 Continuation (2 of 4) of ACN45090 from base 100001 (Human genomic sequence hCC22125.)
 WP Sequence split into 4 fragments LCUS ACN45090 Accession Acn45090
 WP Fragment Name Begin End
 WP ACN45090_0 0 1
 WP ACN45090_-1 100001 210000
 WP ACN45090_-2 200001 310000
 WP ACN45090_-3 300001 350764
 Query Match 11.8%; Score 36.8; DB 11; Length 110000;
 WP Best Local Similarity 56.7%; Pred. No. 2.7;
 Matches 68; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
 Qy 70 TTCAATCATCCAGAAAAGGTGGAAAAAGTTGGATTTGGTCTGAC 129
 Db 76489 TTCTTTTTAAAAAATGATGCAAAGCTGTGCTGGGGTATTACTCGAA 76548
 Qy 130 CATCTCTAGATCGTGGAAAAGCTCACCTTAGCCCTAACGTACATATGGGATTGGAA 189
 Db 76549 ATGTTTAAAGATTGTAATAAGGTGGTGAAGAAACTGAGCTAACATTATGAAATGGAA 76608
 RESULT 5
 ABL02730
 ID ABL02730 standard; cDNA; 7450 BP.
 XX
 AC ABL02730;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 2672.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 XX
 PA (PKB) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-676860/75.
 DR P-PSDB; ABB38627.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more

genes from *Drosophila* and for elucidating cell signalling and cell-cell interactions.

the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/publishedпатsequences.

Claim 1; SEQ ID NO 2672; 21pp + Sequence Listing; English.

Sequence 4074 BP: 1066 A: 930 C: 1005 G: 1070 T: 0 U: 3 Other:

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-AB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp://wipo.int/pub/published/pat_sequences

Query Match	10.6%	Score 33;	DB 10:	Length 4074;
Best Local Similarity	52.6%	Pred. No.	12;	
Matches	72;	Mismatches	65;	Indels 0;
				Gaps 0;
16	AAAGGGCTCGCTAATACCATTTGTTAAAAAGTGTGAA	AAAGTGTGAA	AGATTAC	CATC
3865	AAATGGTCCCCATAATGATAAT	TGGGTC	AAAGCAGCATATTCA	AAAGCACTTT
76	ATCCAGAAAAGGTGTGAAAAGTTGCGATTCGTTGAAA	AAAGTTG	GTCTGACCATC	TC

Sequence	7450 BP;	2044 A;	1752 C;	1603 G;	2051 T;	0 U;	0 Other;
Query Match	11.0%		Score 34.4;	DB 4;	Length 7450;		
Best Local Similarity	57.4%		Pred. No. 5.5;				
Matches	62;	Conservative	0;	Mismatches	46;	Indels	0;
Gaps							

	136	TAGATCGGGAAAAAGT	152
b	3745	GGGATCGTGATACAAAT	3729

71	TCATCATCCGAAAGGTGGAAAAGTGGATTGGCTGGAAAAG 11.8
7128	AAAGAGCGCAGAACCGTTGCAACATGTTTGCCAGTTAAAG 71.75
7068	ATTCTCTTGGCTCAGTAATTCTATTAAGATAAAGGGCAATTAGGGCAATGTGAA 7
11	ATTGAAAGGCCCTGGTAATAACCATTGTCGAAAGTGTAAATCGGAAAGAGAT 7

RESULT 7
 ABK76259 standard; DNA; 468 BP.
 X ABK76259;
 C ABK76259;
 X 13-AUG-2002 (first entry)

RESULT⁶
BSB69083/C
ADD69083 standard; DNA; 4074 BP.

Bacillus licheniformis genomic sequence tag (GST) #3550.
X Differential gene expression; genomic sequenced tag; GST;
WW altered culture condition; environmental stress;

ADB69083 ;
04 -DEC-2003 (First entry)

Bacillus licheniformis.
WO200929113-A2.

S. nelsoniae genomic DNA sequence was 10.33

11-APR-2002.

Cryptococcus neoformans.

WO2003052076-A2.

R R 27-MAR-2001/ 2001US-0279526P.

26-JUN-2003.

A (NOVO) NOVOZYMES BIOTECH INC.

17-DEC-2002: 2002WO-11S040225

INTRODUCTION / 603

17-DEC-2001 : 200111S-0341261B

THE INSTITUTE OF
INDIAN POLITICAL SCIENCE

(BLIT-) ELITRA PHARM INC.
zamudio C, Eroskin AM;
Monitoring differential expression of several genes in first Bacillus
cell relative to expression of same genes in one or more second Bacillus
cells bearing substrate containing Bacillus genomic segment rad

X Monitoring differential expression of several genes in first *Bacillus* cell relative to expression of same genes in one or more second *Bacillus* cells by using substrate containing *Bacillus* genomic segment(s) as

New nucleic acid, useful for preparing a composition for treating an infection.

X array.
X
X Claim 4: SEQ ID NO 3550; 200pp; English.
X
X The invention also includes a method of monitoring differential expression of

CC genes in a first Bacillus cell relative to expression of other Bacillus cells, comprising hybridising labelled nucleic acid probes isolated from Bacillus cells to a substrate containing array of Bacillus genomic sequenced tags (GST), examining the array, and determining relative gene expression by an observed hybridisation reporter signal of a spot in the array. The method is useful for measuring the expression of genes in a first Bacillus cell relative to expression of the same genes in one or more second Bacillus cells. The method is useful for monitoring global expression of several genes from a Bacillus cell, discovering new genes, identifying possible functions of unknown open reading frames and monitoring gene copy number variation and stability. Monitoring changes

CC Claim 3 ; SEQ ID NO 210; 136pp; English.

CC The invention relates to a novel purified or isolated Cryptococcus neoformans nucleic acid molecule comprising a sequence encoding a polypeptide comprising a sequence not given in the specification. A polymucleotide of the invention has fungicidal activity, and may have a use in gene therapy. The nucleic acid is useful for preparing a composition for treating an infection used by Cryptococcus neoformans. The present sequence represents a C. neoformans sequence of the invention. Note: The sequence data for this patent is not represented in

CC in expression of genes may be used to provide a representation of the way
 CC in which Bacillus cells adapt to changes in culture conditions,
 CC environmental stress or other physiological provocation. Extensive follow
 CC -up characterisation is unnecessary when one spot on an array equals one
 CC gene or one open reading frame, since sequence information is available.
 CC This sequence represents a genomic sequence tag (GST) used in the method
 CC of the invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 468 BP; 124 A; 112 C; 114 G; 118 T; 0 U; 0 Other;
 SQ Query Match 10.4%; Score 32.6; DB 6; Length 468;
 Best Local Similarity 55.9%; Pred. No. 7;
 Matches 62; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
 Qy 21 GCCTCGGTTAACATTGTGGAAAAAGTGTGTAATACGAAAGAATTCAATCATCA 80
 Db 191 GTCTCGAAACAGTCGGCTTACACGATCAAATGCCGAAAGGCCATTTTTATCC 250
 Qy 81 GAAAAGGTGTGGAAAAAGTGTGGATTGCCTGGAAAAAGTGTGATCTGACCA 131
 Db 251 GGAAAGGGCTTGAAACTTAGACAGTCATGGAAAGTGGGAGTGTGACAA 301

RESULT 8

ABQ69245_01

Continuation (2 of 31) of ABQ69245 from base 100001 (Listeria innocua DNA sequence #694.

WP Sequence split into 31 fragments LOCUS ABQ69245 Accession Abq69245
 WP Fragment Name Begin End
 WP ABQ69245_00 1 110000
 WP ABQ69245_01 100001 210000
 WP ABQ69245_02 200001 310000
 WP ABQ69245_03 300001 410000
 WP ABQ69245_04 400001 510000
 WP ABQ69245_05 500001 610000
 WP ABQ69245_06 600001 710000
 WP ABQ69245_07 700001 810000
 WP ABQ69245_08 800001 910000
 WP ABQ69245_09 900001 1010000
 WP ABQ69245_10 1000001 1110000
 WP ABQ69245_11 1100001 1210000
 WP ABQ69245_12 1200001 1310000
 WP ABQ69245_13 1300001 1410000
 WP ABQ69245_14 1400001 1510000
 WP ABQ69245_15 1500001 1610000
 WP ABQ69245_16 1600001 1710000
 WP ABQ69245_17 1700001 1810000
 WP ABQ69245_18 1800001 1910000
 WP ABQ69245_19 1900001 2010000
 WP ABQ69245_20 2000001 2110000
 WP ABQ69245_21 2100001 2210000
 WP ABQ69245_22 2200001 2310000
 WP ABQ69245_23 2300001 2410000
 WP ABQ69245_24 2400001 2510000
 WP ABQ69245_25 2500001 2610000
 WP ABQ69245_26 2600001 2710000
 WP ABQ69245_27 2700001 2810000
 WP ABQ69245_28 2800001 2910000
 WP ABQ69245_29 2900001 3010000
 WP ABQ69245_30 3000001 3011208

Query Match 10.4%; Score 32.6; DB 6; Length 110000;

Best Local Similarity 51.7%; Pred. No. 57;

Matches 74; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

Qy 62 AAAGAGATTACATCACAGAAAGTGTGGAAAAAGTGTGGATTAGCGCTTACCTACATATGTTGA 181

Db 46002 AAATGAGGTATAATCAAATC3GCTGAGATGAAGATGAAAGCTT 46061

Qy 122 GATCTGACCATCTAGATCCTGTTACCTACATATGTTGA 181

Db 46062 TTCTAGTAACTCTGTTGTTAGTATCGTTAGTGTGTTCAATTGTTGA 46121

CC 182 TTGGGAAAAAGAGACGGAGGC 204
 CC 46122 GAGTCAAACGGGAAACGGGC 46144

RESULT 9
 ABQ67197_00
 WP Sequence split into 12 fragments LOCUS ABQ67197 Accession Abq67197

WP Fragment Name Begin End
 WP ABQ67197_00 1 110000
 WP ABQ67197_01 100001 210000
 WP ABQ67197_02 200001 310000
 WP ABQ67197_03 300001 410000
 WP ABQ67197_04 400001 510000
 WP ABQ67197_05 500001 610000
 WP ABQ67197_06 600001 710000
 WP ABQ67197_07 700001 810000
 WP ABQ67197_08 800001 910000
 WP ABQ67197_09 900001 1010000
 WP ABQ67197_10 1000001 1110000
 ID ABQ67197_11 1100001 1163020 BP.
 XX ABC67197;
 AC AC ABC67197;
 XX DT 29-AUG-2002 (first entry)

XX Listeria innocua contig DNA sequence #10.
 DE DB Listeria innocua contig DNA sequence #10.

WP Antibacterial; Listeria; food contamination; mutational analysis;
 KW KW infection; ds.
 XX OS Listeria innocua.
 XX PN WO200228891-A2.
 XX PD 11-APR-2002.
 XX PF 04-OCT-2001; 2001WO-PR003061.
 XX PR 04-OCT-2000; 2000FR-00012697.
 XX PA (INSP) INST PASTEUR.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX PR treatment and prevention of infection, also related polypeptides,
 PT Kunst F, Glasner P;
 XX DR WPI; 2002-332479/37.
 XX New genomic sequences from Listeria species, useful for detection,
 PT treatment and prevention of infection, also related polypeptides,
 PT antibodies and modulators.

XX Claim 5; SEQ ID NO 10; 180pp; French.
 XX The present invention relates to nucleic acid sequences (ABQ67197-
 CC ABQ71212) from Listeria sp. The sequences are useful as probes and
 CC primers for identification and/or detection of Listeria (e.g. as
 CC contaminants in foods, or mutational analysis) and for analysis of gene
 CC expression. Proteins encoded by the nucleic acid sequences can be used to
 CC screen for compounds that modulate gene expression, replication and
 CC pathogenicity of Listeria (potential therapeutic agents), also for
 CC treating infections by Listeria, and are useful as immunogens in anti-
 CC Listeria vaccines. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1163020 BP; 388339A; 197589C; 235612S; 341474T; 0U; 60ther;
 CC Query Match 10.4%; Score 32.6; DB 6; Length 110000;
 CC Best Local Similarity 51.7%; Pred. No. 57;
 CC Matches 74; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

CC SQ Sequence Match 10.4%; Score 32.6; DB 6; Length 110000;
 CC Best Local Similarity 51.7%; Pred. No. 57;
 CC Matches 74; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

This sequence encodes a Streptococcus pneumoniae protein that (based on homology with a *Bacillus subtilis* protein) is a BSCEIABCD NCBI gi: CC 895746, and represents a DNA sequence of the invention. The DNA sequences were isolated from *Streptococcus pneumoniae* strain 010993 (NCIMB 40794). The *Streptococcus pneumoniae* proteins of the invention can be used to identify compounds which interact with and inhibit or activate the activity of the proteins. Antagonists can be used to treat diseases caused by *S. pneumoniae* proteins, through genetic immunization. They can also be used to induce an immunological response in a mammal by inoculation with the *S. pneumoniae* proteins or delivery of the encoding nucleic acids in a vector adequate to produce antibody and/or T cell immune responses to protect the animal from disease. The proteins can also be used to identify antimicrobial compounds which are capable of inhibiting their bioactivity. In particular the proteins of the invention can be used to prevent adhesion of bacteria to mammalian extracellular matrix proteins on in-dwelling devices or in wounds, to block protein-mediated mammalian cell invasion, and to block the normal progression of pathogenesis in infections initiated other than by the implantation of in-dwelling devices or other surgical techniques

XX Sequence 1096 BP; 338 A; 224 C; 138 G; 396 T; 0 U; 0 Other;

SQ Query Match 10.4%; Score 32.4%; Pred. No. 11; DB 2; Length 1096;

Best Local Similarity 51.4%; Mismatches 0; Indels 71; Gaps 0; Matches 75; Conservative 0; MisMatches 0; Del 0; Gaps 0;

QY 62 AAAGGAGATTCACTCATCCAGAAAAGGGCTGGAAAAAGTTGCGTGGAAAAAGTC 121
DB 699 AGACAGGTAAATTACTTGAAAGATGGTGAAGTTATCCTTGTTGAATCAAATGA 640QY 122 GATCTGACCATCTTAGATCTGTTGAAAAGTTACGTAGCTTACGTACATATGTCGA 181
DB 639 TTCTGAAACAACTCTTACTTTATGAAATTAAACAGCGCCAGAATTAACTCAGTTGTTT 580QY 182 TTGTTGAAAGAAAGAACGGAGGCATC 207
DB 579 TTCTGAAAAATTCAGTCCGCACT 554RESULT 12
AA296371/C

ID AA296371 standard; DNA; 1727 BP.

XX DT 10-APR-2000 (first entry)

S. pneumoniae derived DNA from ORF #199.

DE Treatment; prevention; disease; diagnosis; gene therapy; screening; KW bacterial; antimicrobial; antibiotic; pathogenesis; infection; OS Streptococcus pneumoniae.
XX PN WO9806734-A1.
XX PD 19-FEB-1998.
XX PF 15-AUG-1997; 97WO-US014436.
XX PR 16-AUG-1996; 96US-0024022P.
XX PA (SMIK) SMITHKLINE BEECHAM CORP.
XX PI Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO;
XX PI Stodola RK;
XX DR WPI; 1998-159452/14.
XX P-PSDB; AAY86040.

PT Streptococcus pneumoniae proteins and related DNA - useful for screening compounds for antibacterial activity.

XX PS Claim 4; Page 230-231; 640pp; English.
XX PS Claim 4; Page 230-231; 640pp; English.
XX CC This invention describes novel isolated *Streptococcus pneumoniae* proteins (see CC AA95732-Y86182). The DNA, vectors and host cells described in the method CC of the invention are useful for the recombinant expression of the CC polypeptides. The polypeptides are useful for treatment or prevention of CC disease, or diagnosis of disease related to expression or activity of CC such a polypeptide. They can also be used to screen for compounds which CC interact with and inhibit or activate such a polypeptide. The CC polypeptides (or DNA encoding them, via gene therapy) are also useful for CC inducing an immunological response in a mammal. The antagonists are CC useful to inhibit such bacterial polypeptides. The polypeptides are CC particularly useful to identify antimicrobial compounds and antibiotics. CC They are also useful to determine their role in pathogenesis of CC infection, dysfunction and disease.

XX SQ Sequence 1727 BP; 546 A; 332 C; 221 G; 628 T; 0 U; 0 Other;

XX SQ Query Match 10.4%; Pred. No. 13; DB 2; Length 1727;

Best Local Similarity 51.4%; Mismatches 0; Indels 71; Gaps 0; Gaps 0; Matches 75; Conservative 0; MisMatches 0; Del 0; Gaps 0;

QY 62 AAAGGAGATTCACTCATCCAGAAAAGGGCTGGAAAAAGTTGCGTGGAAAAAGTC 121
DB 696 AGCAAGGTAAATTACTTGAAAGATGGTGAAGTTATCCTTGTTGAATCAAATGA 637QY 122 GATCTGACCATCTTAGATCTGTTGAAAAGTTACGTAGCTTACGTACATATGTCGA 181
DB 636 GTTGTGAACTCTTACTTTATGAAATTAAACAGCGCAAGATAATGTTCACTGTTT 577QY 182 TTGTTGAAAGAAAGAACGGAGGCATC 207
DB 576 TTCTGAAAAATTCAGTCCGCACT 554RESULT 13
AA166078

ID AA166078 standard; DNA; 2992 BP.

XX AC AA166078;

XX DT 14-JAN-2002 (first entry)

XX DE Spinacia oleracea ClpC protease gene (GenBank: AF043539).

XX KW Transgenic plant; stress tolerance; heat shock protein; HSP; cotton;

KW canola; soybean; corn; wheat; tobacco; sorghum; potato; tomato;

KW Arabidopsis thaliana; db.

XX OS Spinacia oleracea.

XX PN WO2001070929-A2.

XX PD 27-SEP-2001.

XX PF 20-MAR-2001; 2001WO-US008336.

XX PR 20-MAR-2000; 2000US-0190769P.

XX PR 18-APR-2000; 2000US-0198116P.

XX (ARCH-) ARCH DEV CORP.

XX PI Lindquist S, Queitsch C, Vierling E;

XX DR WPI; 2001-639123/73.

XX PT Transgenic plants with improved heat stress tolerance, useful for producing animal feed, oil and synthetic products.

XX PS Claim 4; Page; 91pp; English.

XX PT

The invention relates to a transgenic plant, comprising a genetic construct comprising a promoter operatively linked to a nucleic acid sequence (AAI605-AAI6084) encoding a plant Heat Shock Protein (HSP) family amino acid sequence (AAMS1651-AAMS1671). The transgenic plant has increased stress tolerance, especially to heat. The plant is a cereal, grass, ornamental plant, crop plant, food plant, oil-producing plant, a synthetic product-producing plant, an environmental waste absorbing plant, an alcohol plant, a medicinal plant, a recreational plant and/or an animal feed plant. In particular, the transgenic plant is cotton, canola, soybean, corn, wheat, tobacco, sorghum, potato, tomato and/or Arabidopsis thaliana. The plants may be used to produce animal feed, alcohol, crop, oil, medicine or a synthetic product. Note: The sequence data for this patent did not form part of the printed specification, but was obtained from GenBank using the Accession Number reference provided in the specification.

Sequence 2992 BP; 847 A; 554 C; 777 G; 814 T; 0 U; 0 Other;

Query Match Score 32.4; DB 4; Length 2992;
Best Local Similarity 54.1%; Pred. No. 16;
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Qy 82 AAAGGTGTGGAAGATGTGTGGATTGCCGTTACGATCTGACCATCTAGATC 141
Db 1176 AAATACCGTGGAGATTAGTGGAAAGATGTAAGAGTGATGGAGGAATTACAGT 1235

Qy 142 GTGAAATAAGTCAGTGTAGCTTACGACATATGCAATGTTGAAAGATGATGGAGGA 201
Db 1236 GATGAAATAATTATTATTATTGATGAAAGTACATCTGATTGGCCGGAGCAGAA 1295

Qy 202 GG 203
Db 1296 GG 1297

RESULT 14
ABN80109 standard; DNA; 6380 BP.
ID ABN80109;
AC XX
XX DR 15-JUL-2002 (first entry)
DE Human chemically modified disease associated gene SEQ ID NO 126.
XX KW Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;
KW heart disease; epilepsy; histone deacetylation; muscular dystrophy;
KW dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;
KW antidiabetic; cytostatic; anticonvulsant; ds.
XX OS Homo sapiens.
OS Synthet.c.
XX PN WO200200927-A2.
XX PD 03-JAN-2002.
XX PP 02-JUL-2001; 2001WO-EP007536.
XX PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX PA (EPIC)- EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K;
XX DR WPI; 2002-130908/17.
XX PR Novel nucleic acid useful for diagnosis and therapy of diseases
PR associated with development genes such as diabetes, comprises a sequence
PR of a segment of chemically pretreated DNA of genes associated with
PR development.

PS Claim 1; SEQ ID NO 126; 27pp; English.
XX The invention relates to a nucleic acid (I) comprising a sequence at least 18 bases in length of a segment of chemically pretreated DNA (II) of genes associated with development selected from 87 genes listed in the specification such as ACCN, ADFN, or AFDI and comprising one of 350 sequences (ABN7984-ABN8033) or their complements. The invention is useful for the diagnosis or therapy of diseases associated with development genes, in particular disease related to homeobox containing genes (HOX), like diabetes, cancer, apoptosis, syndromes associated with congenital heart disease, epilepsy, diseases related to histone deacetylation, Curarino syndrome, diseases related with the development of the brain and limb girdle muscular dystrophy and dwarfism. Oligomers specific to each of the genes are useful for detecting the methylation state of all CpG dinucleotides within the 350 sequences (II) and their complementary sequences, as primer oligonucleotides for the amplification of the 350 sequences (II) and/or their complements and as oligomer probes for detecting the cytosine methylation state and/or single nucleotide polymorphisms (SNPs). Note: The sequence data for this patent did not form part of the printed specification, but is based on sequence information supplied to the European Patent Office

Sequence 6380 BP; 1663 A; 193 C; 1758 G; 2765 T; 0 U; 1 Other;

SQ Query Match Score 32.4; DB 6; Length 6380;

SQ Best Local Similarity 62.2%; Pred. No. 22;

SQ Matches 51; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

SQ 46 ATTTGGATAATCGGAAAAGAGATTCATCATCCAGAAAAGGTGTGGAAAGTTGTGGAT 105
Db 5808 ATTTGATAAAAAGGAAAATAGTTTAAATAGGAATAAGGTAAATAGGAAT 5867

SQ 106 TGGCTGGAAAAGACTCGATCTG 127
Db 5868 TCGTAGATAGGTAAATGTG 5889

RESULT 15

ABV52248/C
ID ABV52248 standard; DNA; 11303 BP.

XX AC AAV52248;
XX DT 23-OCT-1998 (first entry)

XX DE Streptococcus pneumoniae genome fragment SEQ ID NO:115.

XX Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;

KW computer readable medium; vaccine; pharmaceutical composition;

KW Streptococcus pneumoniae.

XX OS WO9818931-A2.
XX PN WO9818931-A2.
XX PD 07-MAY-1998.
XX PF 30-OCT-1997; 97WO-US019588.
XX PR 31-OCT-1996; 96US-0029960P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Kunisch CA, Choi GH, Dillon PJ, Rosen CA, Barash SC, Fannon M;
PI Dougherty BA;
XX DR WPI; 1998-272225/24.

XX Computer-readable medium with recorded Streptococcus pneumoniae polynucleotide sequences - useful in diagnostic kits and assays, and pharmaceutical compositions and vaccines for Streptococcus pneumoniae.
XX Claim 1; Page 831-838; 1409pp; English.

The present invention describes a computer readable medium which has the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52524) recorded on it, or a representative fragment or a sequence at least 95% identical to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in SEQ ID NO:1 to 391 (AAV52134 to AAV52524) are genomic fragments from *Streptococcus pneumoniae*. The present invention also describes an isolated nucleic acid molecule encoding a homologue of any of the fragments of the *S. pneumoniae* genome (SEQ ID NO:1 to 391) where the nucleic acid molecule is produced by a process comprising: (a) screening a genomic DNA library using as a probe a target sequence defined by any of the sequences in SEQ ID NO:1 to 391, identifying members of the library which contain sequences that hybridise to the target sequence and isolating the nucleic acid molecules from the members; or (b) isolating mRNA, DNA or cDNA produced from an organism, amplifying nucleic acid molecules whose nucleotide sequence is homologous to amplification primers derived from the fragment of the *S. pneumoniae* genome to prime the amplification and isolating the amplified sequences. The computer readable medium can be used in a computer-based system for identifying fragments of the *S. pneumoniae* genome of commercial importance, or expression modulating fragments of the *S. pneumoniae* genome. Products from the present invention can be used in diagnosis kits and assays, and pharmaceutical compositions and vaccines for *S. pneumoniae*

XX SQ Sequence 11303 BP; 3733 A; 2266 C; 1558 G; 3746 T; 0 U; 0 Other;

	Query Match	Score	DB	Length
Qy	10.4%; Best Local Similarity Matches 75; Conservative 0; Mismatches 71; Indels 0; Gaps 0;	32.4;	2;	11303;
Db	62 AAAGAGATTCATCATCAGAAAAACGTGTTGAAAGTGTGGATTGCGTGGAAAAGTTC 121 6373 AGACGAGTAATTTACTTGAAAGATGGTGAAGTTTCAATGAACTTAAATGGA 6314			
Qy	122 GATCTGACCATCTCTAGATGTTGAAAAAAGTTCAAGTACGTAGCTTACCTAACATATGSGA 181			
Db	6313 GTTTGAACTCTTACTTTATGGATAAACAGGGCAAGAATTAAACAGGGCAAGTCAATTGTT 6254			
Qy	182 TTGTGAAAGAAAGAACGAGGGCATC 207			
Db	6253 TTCTGAAAGAAATTGGAGTCGCATC 6228			

Search completed: July 3, 2005, 02:38:02
Job time : 481.353 secs

This Page Blank (USP1)

Result No.	Score	Query Match	Length DB	ID	Description
C 1	37.4	12.0	2515	13	US-10-027-632-59048
C 2	37.4	12.0	2515	19	Sequence 59048, A
C 3	36.8	11.8	350764	13	Sequence 59048, A
C 4	33.8	10.6	1491	19	Sequence 59048, A
C 5	33.3	10.6	4074	17	Sequence 59048, A
C 6	32.8	10.5	657	13	Sequence 59048, A
C 7	32.8	10.5	657	17	Sequence 59048, A

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Match	Length	DB ID
C 8	9	32.8	10.5	145806	20 US-10-719-993-6943
C 10	9	32.6	10.4	468	19 US-10-974-300-3550
C 11	10	32.6	10.4	585	19 Sequence 29171, A
C 12	11	32.6	10.4	1163020	17 US-10-398-221-10
C 13	12	32.6	10.4	3011208	17 US-10-398-221-10
C 14	13	32.4	10.4	495	21 US-10-472-928-445
C 15	14	32.4	10.4	2992	9 US-09-812-350-43
C 16	15	32.4	10.4	11303	8 US-08-961-527-115
C 17	16	32.4	10.4	11303	17 US-10-158-844-115
C 18	17	32.4	10.4	65454	18 Sequence 115, App
C 19	18	32.4	10.4	83517	19 Sequence 11, App
C 20	19	32.4	10.4	103660	21 US-10-741-600-17645
C 21	20	32.4	10.4	2162598	21 Sequence 47979, App
C 22	21	32.2	10.3	370	19 US-10-472-928-4979
C 23	22	32.2	10.3	625	13 Sequence 4790, App
C 24	23	32.2	10.3	625	17 Sequence 236927,
C 25	24	32.2	10.3	625	17 Sequence 236928,
C 26	25	32	10.3	83517	19 Sequence 236927,
C 27	26	32	10.3	765	13 Sequence 236928,
C 28	27	32	10.3	765	13 Sequence 236927,
C 29	28	32	10.3	765	13 Sequence 236927,
C 30	29	32	10.3	765	17 Sequence 236927,
C 31	30	31.8	10.2	722	17 Sequence 236927,
C 32	31	31.8	10.2	2222	17 Sequence 236927,
C 33	32	31.6	10.1	592	9 Sequence 236927,
C 34	33	31.6	10.1	2526	10 Sequence 1116, App
C 35	34	31.6	10.1	2526	15 Sequence 1116, App
C 36	35	31.6	10.1	2526	17 Sequence 1116, App
C 37	36	31.6	10.1	2526	17 Sequence 1116, App
C 38	37	31.6	10.1	2526	18 Sequence 1116, App
C 39	38	31.6	10.1	2526	18 Sequence 1116, App
C 40	39	31.6	10.1	2526	18 Sequence 1116, App
C 41	40	31.4	10.1	443	13 Sequence 1116, App
C 42	41	31.4	10.1	443	17 Sequence 1116, App
C 43	42	31.4	10.1	367378	16 Sequence 1, Appli
C 44	43	31.2	10.0	9859	15 Sequence 2105, AP
C 45	44	31	9.9	783	18 Sequence 97972, A

ALIGNMENTS

Scoring table: IDENTITY_NUC Gapop 10.0 , Gapext 1.0
 Perfect score: 312
 Sequence: 1 gtcgaccatatt-gaaagg.....tgaaccttgcgtact 312
 Title: US-10-814-858A-1
 Run on: July 3, 2005, 03:19:18 ; Search time 640.235 Seconds
 (without alignments)
 3056.566 Million cell updates/sec

Scoring table: IDENTITY_NUC Gapop 10.0 , Gapext 1.0
 Searched: 6313374 seqs., 3136092125 residues
 Total number of hits satisfying chosen parameters: 12626748
 Minimum DB seq length: 0
 Maximum DB seq length: 20000000000
 Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Published Applications_NA.*
 1: /cgcn2_6/ptodata/1/pubpna/us007_PUBCOMB.seq/*
 2: /cgpn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq/*
 3: /cgpn2_6/ptodata/1/pubpna/US006_PUB.seq/*
 4: /cgpn2_6/ptodata/1/pubpna/us007_PUBCOMB.seq/*
 5: /cgpn2_6/ptodata/1/pubpna/us007_NEW_PUB.seq/*
 6: /cgpn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq/*
 7: /cgpn2_6/ptodata/1/pubpna/us008_NEW_PUB.seq/*
 8: /cgpn2_6/ptodata/1/pubpna/us008_PUBCOMB.seq/*
 9: /cgpn2_6/ptodata/1/pubpna/us009_PUBCOMB.seq/*
 10: /cgpn2_6/ptodata/1/pubpna/us009_NEW_PUB.seq/*
 11: /cgpn2_6/ptodata/1/pubpna/us010_PUBCOMB.seq/*
 12: /cgpn2_6/ptodata/1/pubpna/us010_NEW_PUB.seq/*
 13: /cgpn2_6/ptodata/1/pubpna/us010_PUBCOMB.seq/*
 14: /cgpn2_6/ptodata/1/pubpna/us010_PUBCOMB.seq/*
 15: /cgpn2_6/ptodata/1/pubpna/us010_PUBCOMB.seq/*
 16: /cgpn2_6/ptodata/1/pubpna/us010D_PUBCOMB.seq/*
 17: /cgpn2_6/ptodata/1/pubpna/us010E_PUBCOMB.seq/*
 18: /cgpn2_6/ptodata/1/pubpna/us010F_PUBCOMB.seq/*
 19: /cgpn2_6/ptodata/1/pubpna/us010G_PUBCOMB.seq/*
 20: /cgpn2_6/ptodata/1/pubpna/us010H_PUBCOMB.seq/*
 21: /cgpn2_6/ptodata/1/pubpna/us010I_PUBCOMB.seq/*
 22: /cgpn2_6/ptodata/1/pubpna/us010_NEW_PUB.seq/*
 23: /cgpn2_6/ptodata/1/pubpna/us11A_PUBCOMB.seq/*
 24: /cgpn2_6/ptodata/1/pubpna/us11I_NEW_PUB.seq/*
 25: /cgpn2_6/ptodata/1/pubpna/us600_NEW_PUB.seq/*
 26: /cgpn2_6/ptodata/1/pubpna/us600_PUBCOMB.seq/*

RESULT 1
 US-10-027-632-59048/c
 Sequence 59048, Application US/10027632
 Publication No. US2002019837IA1
 GENERAL INFORMATION:
 APPLICANT: Wang, David G.
 TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome

TITLE OF INVENTION: Polymorphisms in the Human Genome

FILE REFERENCE: 108827-129
 CURRENT APPLICATION NUMBER: US-10-027-632-59048
 CURRENT FILING DATE: 2002-04-30
 PRIOR APPLICATION NUMBER: US 60/185,218
 PRIOR FILING DATE: 2000-07-12
 PRIOR APPLICATION NUMBER: US 60/198,676
 PRIOR FILING DATE: 2000-04-20
 PRIOR APPLICATION NUMBER: US 60/193,483
 PRIOR FILING DATE: 2000-03-29
 NUMBER OF SEQ ID NOS: 325720
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 59048
 LENGTH: 2515
 TYPE: DNA

QY 41 GAAAAGTTGTAATACGAAAAAGAAGATTCACTCCAGAAAAGGCTGGAAAAGCTG 100
 Db 1126 GAAATTGTCGAAATTGCAAAAGAATTCTTGTGAAAGAAATTCTTGGAAAAGTT 1185
 QY 101 TGGATTCGCGGAAAGGTTGACATCTGACCATCTGATCGTGAAGAAAGTTCACTGTA 160
 Db 1186 TCCGATGTCGAAAGAACCTCAACATGATTCCTGATCACTAAAGCAAAATCAGCTTT 1245
 QY 161 GCGCTTACGPACA 173
 Db 1246 GGTCGAGAACCA 1258

RESULT 5
 US-10-320-797-210/c
 i Sequence 210, Application US/10320797
 i GENERAL INFORMATION:
 i APPLICANT: Eroshkin, Alexey M.
 i TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS NEOFORMANS AND
 i TITLE OF INVENTION: METHODS OF USE
 i FILE REFERENCE: 10102-021-999
 i CURRENT APPLICATION NUMBER: US/10/320,797
 i CURRENT FILING DATE: 2002-12-16
 i PRIOR APPLICATION NUMBER: 60/341,261
 i PRIOR FILING DATE: 2001-12-17
 i NUMBER OF SEQ ID NOS: 3361
 i SOFTWARE: PatentIn version 3.1
 i SEQ ID NO: 210
 i LENGTH: 4074
 i TYPE: DNA
 i ORGANISM: Cryptococcus neoformans

US-10-320-797-210

Query Match 10.6%; Score 33; DB 17; Length 4074;
 Best Local Similarity 52.6%; Pred. No. 23;
 Matches 72; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
 QY 16 AAAGGGCTCTGGTATAACCATTTGGAAAGGTTAACCGAAAAGAAAGATTCTC 75
 Db 3865 AAATGGTCCCATAATGATAATGGTCAAGGAGCATATTCAAAAGAACCTTT 3806
 QY 76 ATCCGAAANAGGTTGCAAAGTGTGCTGGAAAAGTGTGATCTGCCATCTC 135
 Db 3805 TGTAATAAAGGATGAGATGTTGAGCTTCTAAAGATTGAACTATGG 3746
 QY 136 TAGATCGTGCAAAGT 152
 Db 3745 GGGATGGTGATCAAT 3729

RESULT 6

US-10-007-632-259460/c
 i Sequence 259460, Application US/10027632
 i GENERAL INFORMATION:
 i APPLICANT: Wang, David G.
 i TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 i TITLE OF INVENTION: Polymorphisms in the Human Genome
 i FILE REFERENCE: 108827-129
 i CURRENT APPLICATION NUMBER: US/10/027,632
 i CURRENT FILING DATE: 2002-04-30
 i PRIOR APPLICATION NUMBER: US 60/218,006
 i PRIOR FILING DATE: 2000-07-12
 i PRIOR APPLICATION NUMBER: US 60/198,676
 i PRIOR FILING DATE: 2000-04-20
 i PRIOR APPLICATION NUMBER: US 60/193,483
 i PRIOR FILING DATE: 2000-03-29
 i PRIOR APPLICATION NUMBER: US 60/185,218
 i PRIOR FILING DATE: 2000-02-24
 i PRIOR APPLICATION NUMBER: US 60/167,363
 i PRIOR FILING DATE: 1999-11-23
 i PRIOR APPLICATION NUMBER: US 60/156,358
 i PRIOR FILING DATE: 1999-09-28
 i PRIOR APPLICATION NUMBER: US 60/146,002
 i NUMBER OF SEQ ID NOS: 325720
 i SOFTWARE: Fast-SEQ for Windows Version 4.0
 i SEQ ID NO: 259460
 i LENGTH: 657
 i TYPE: DNA
 i ORGANISM: Human

US-10-027-632-259460/c
 i Sequence 259460, Application US/10027632
 i GENERAL INFORMATION:
 i APPLICANT: Wang, David G.
 i TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 i TITLE OF INVENTION: Polymorphisms in the Human Genome
 i FILE REFERENCE: 108827-129
 i CURRENT APPLICATION NUMBER: US/10/027,632
 i CURRENT FILING DATE: 2002-04-30
 i PRIOR APPLICATION NUMBER: US 60/218,006
 i PRIOR FILING DATE: 2000-04-20
 i PRIOR APPLICATION NUMBER: US 60/198,676
 i PRIOR FILING DATE: 2000-03-24
 i PRIOR APPLICATION NUMBER: US 60/185,218
 i PRIOR FILING DATE: 1999-11-23
 i PRIOR APPLICATION NUMBER: US 60/156,358
 i PRIOR FILING DATE: 1999-09-28
 i PRIOR APPLICATION NUMBER: US 60/146,002
 i NUMBER OF SEQ ID NOS: 325720
 i SOFTWARE: Fast-SEQ for Windows Version 4.0
 i SEQ ID NO: 259460
 i LENGTH: 657
 i TYPE: DNA
 i ORGANISM: Human

RESULT 7
 US-10-027-632-259460/c
 i Sequence 259460, Application US/10027632
 i GENERAL INFORMATION:
 i APPLICANT: Wang, David G.
 i TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 i TITLE OF INVENTION: Polymorphisms in the Human Genome
 i FILE REFERENCE: 108827-129
 i CURRENT APPLICATION NUMBER: US/10/027,632
 i CURRENT FILING DATE: 2002-04-30
 i PRIOR APPLICATION NUMBER: US 60/218,006
 i PRIOR FILING DATE: 2000-07-12
 i PRIOR APPLICATION NUMBER: US 60/198,676
 i PRIOR FILING DATE: 2000-04-20
 i PRIOR APPLICATION NUMBER: US 60/193,483
 i PRIOR FILING DATE: 2000-03-24
 i PRIOR APPLICATION NUMBER: US 60/185,218
 i PRIOR FILING DATE: 1999-11-23
 i PRIOR APPLICATION NUMBER: US 60/156,358
 i PRIOR FILING DATE: 1999-09-28
 i PRIOR APPLICATION NUMBER: US 60/146,002
 i NUMBER OF SEQ ID NOS: 325720
 i SOFTWARE: Fast-SEQ for Windows Version 4.0
 i SEQ ID NO: 259460
 i LENGTH: 657
 i TYPE: DNA
 i ORGANISM: Human

US-10-719-993-6943

```

Sequence 6943, Application US/10719993
; PUBLICATION NO. US2004025849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michael et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01496
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 6943
; LENGTH: 145816
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-19-93-6943

Query Match 10.5%; Score 32.8; DB 20; Length 145806;
Best Local Similarity 53.0%; Pred. No. 1.1e+02; Indels 0; Gaps 0;
Matches 70; Conservative 0; Mismatches 62; Gaps 0;
Qy 8 ATCATTGAAAGGCCCTCGGTAAATACATTGTCGAAAAGTTGCTTAATACGGAAAAAGRA 67
Db 82601 ATTCACTGCAGGTTGTTAAAGGCTGAAAGCTGAAARGCCAANAGGGAATACAGGTTCAAA 83660
Qy 68 GATTCATCATCCGAAAGGTGTGAAAGGTGTGGATTGTGGTGAAGAAGTTCGATCTG 127
Db 82661 GATAAATTCCAGAAAGGTGTATGGCTCTGGATGTAGAGCAAAAGAAATAA 82720
Qy 128 ACCATCTCTAGA 139
Db 82721 ATATTACCCAGA 82732

RESULT 9
US-09-974-300-3550
; Sequence 3550, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene Expression
; FILE REFERENCE: 10085_500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 3550
; LENGTH: 468
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-3550

Query Match 10.4%; Score 32.6; DB 9; Length 468;
Best Local Similarity 55.9%; Pred. No. 1.2e+02; Indels 0; Gaps 0;
Matches 62; Conservative 0; Mismatches 49; Gaps 0;
Qy 21 GCCTCGTAAATACATTGTCGAAAAGTTGGTGTACCTGGTAAATACGGCTCATCCOA 80
Db 191 GTCCTGAACTCGGCTTACAGCTCAAAATGGCGGAATAGGGCATTTTATCC 250
Qy 81 GAAGAGGTGTGAAAAGTTGTGGATTCGGTGGAAAAGTTGATCTGACCA 131
Db 251 GGAAAGGGCTTGAAACTTAGACAGTCATGGAAAATGGGGCTGTGCAA 301

RESULT 10
US-10-7167-701-2310
; Sequence 29171, Application US/10767701
; PUBLICATION NO. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(5353)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO: 29171
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE: OTHER INFORMATION: Clone ID: 9296737
; US-10-767-701-29171

Query Match 10.4%; Score 32.6; DB 19; Length 585;
Best Local Similarity 55.9%; Pred. No. 1.3e+02; Indels 0; Gaps 0;
Matches 62; Conservative 0; Mismatches 49; Gaps 0;
Qy 88 TGTGAAAGAAGTTGTGGATTCGGTAAAGAAGTTGATCTGACATCTGATCTGCGAA 147
Db 102 TGATCAAGGTATAGTGTGGTTGAAACCGCGGCAAACCAGGGCTGACTTCGCGAA 161
Qy 148 AAAAGTTCACCGTTAGCCCTAGGTACATATGTGGATTTGTGGAAAAGAGAC 198
Db 162 AAGCTCGACCTGAGGCATTCGGTAAAGCTATGCGGAGGTGCAAGGAGGCC 212

RESULT 11
US-10-398-221-10
; Sequence 10, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Friderik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344_702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 10
; LENGTH: 1163020
; TYPE: DNA
; ORGANISM: Listeria innocua
; FEATURE: misc_feature
; NAME/KEY: (1)..(end)
; OTHER INFORMATION: n can be any nucleotide: a, g, c or t/u
; US-10-398-221-10

Query Match 10.4%; Score 32.6; DB 17; Length 1163020;
Best Local Similarity 51.7%; Pred. No. 3.6e+02; Indels 0; Gaps 0;
Matches 74; Conservative 0; Mismatches 69; Gaps 0;
Qy 62 AAAGAAAGATTCATCATCCAGAAAAGTTGTGGATTCGGTGGAAAAAGCTTC 121
Db 61761 AAATAGGGTTATAATCATAAATCGGGTAAAGATGAAATAAGAGATGAAAGGGTT 61820
Qy 122 GATCTGACCATCTCTGATCTGGCTTACGGTACAGTTCAAGTTCGCGA 181
Db 61821 TTATGATTTACTTCCTGTTGTTGTGTTGATTTATCCATTGCTGTGTC 61880

RESULT 12
US-10-7167-701-2310
; Sequence 29172, Application US/10767702
; PUBLICATION NO. US20040172685A1
; GENERAL INFORMATION:
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(5353)B
; CURRENT APPLICATION NUMBER: US/10/767,702
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63129
; SEQ ID NO: 29172
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE: OTHER INFORMATION: Clone ID: 9296738
; US-10-767-701-29172

Query Match 10.4%; Score 32.6; DB 17; Length 585;
Best Local Similarity 55.9%; Pred. No. 1.3e+02; Indels 0; Gaps 0;
Matches 62; Conservative 0; Mismatches 49; Gaps 0;
Qy 88 TGTGAAAGAAGTTGTGGATTCGGTAAAGAAGTTGATCTGACATCTGATCTGCGAA 147
Db 102 TGATCAAGGTATAGTGTGGTTGAAACCGCGGCAAACCAGGGCTGACTTCGCGAA 161
Qy 148 AAAAGTTCACCGTTAGCCCTAGGTACATATGTGGATTTGTGGAAAAGAGAC 198
Db 162 AAGCTCGACCTGAGGCATTCGGTAAAGCTATGCGGAGGTGCAAGGAGGCC 212

```

RESULT 12
US-10-398-221-2058
; Sequence 2058, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIORITY APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 2058
; TYPE: DNA
; ORGANISM: Listeria innocua

US-10-398-221-2058
Query Match Similarity 10.4%; Score 32.6; DB 17; Length 3011208;
Matches 74; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
Query 62 AAAGAAGATTCAATCCGAAAGGTGTGAAAGATGTGGATTGGCTGGAAAAAGTC 121
Db 146002 AAATGAGGTTATAATCAAATCGGGTGAAGATAAGATGAAAAGAGTT 146061
Query 122 GATCTGACCTCTAGATCCTGCGAAANAGTTAACGTTAGGCCATTACATATGCGA 181
Db 146062 TTATGTTATCATCTGTGTGTGTAGTTATCATGTTAGTGTGTCAATTGGTGC 146121
Query 182 TGTGAAAGAGAGACGGGGC 204
Db 146122 GACTCAANACGCGCGAAGCGGC 146144

RESULT 13
US-10-472-928-445
; Sequence 445, Application US/10472928
; Publication No. US20050020813A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STREPTOCOCUS PNEUMONIAE PROTEINS AND NUCLEAR ACIDS
; FILE REFERENCE: P026926HO
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIORITY APPLICATION NUMBER: GB-01-07658.7
; PRIORITY FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO: 445
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae

US-10-472-928-445
Query Match Similarity 10.4%; Score 32.4; DB 21; Length 495;
Matches 75; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
Query 122 GATCTGACCTCTAGATCCTGCGAAAGGTGTGGATTGGCTGGAAAAAGTC 121
Db 165 GTTGAAACATCTTACTTTATGGAAATTAAAGCGCGAATTAAGTTCAGTT 250
Query 182 TTGTGAAAGAGAGACGGGGC 204
Db 225 TTCTGAAAAAATTCGTTCCGCATC 250

RESULT 14
US-10-398-221-2059
; Sequence 2059, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: Lindquist, Susan
; APPLICANT: Quitsch, Christine
; APPLICANT: Vierling, Elizabeth
; TITLE OF INVENTION: Transgenic Plants Containing Heat Shock Protein
; FILE REFERENCE: P01979US2
; CURRENT APPLICATION NUMBER: US/09/812,350
; CURRENT FILING DATE: 2001-03-20
; PRIORITY APPLICATION NUMBER: US 60/190,769
; PRIORITY FILING DATE: 2000-03-20
; PRIORITY APPLICATION NUMBER: US 60/198,116
; SEQ ID NO: 43
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 43
; TYPE: DNA
; ORGANISM: Spinacia oleracea

US-09-812-350-43
Query Match Similarity 10.4%; Score 32.4; DB 9; Length 2992;
Best Local Similarity 54.1%; Pred. No. 31; Mismatches 56; Indels 0; Gaps 0;
Matches 66; Conservative 0; Number of SEQ ID NOS: 49
; TYPE: DNA
; ORGANISM: Spinacia oleracea

US-09-812-350-43
Query 82 AAAGGTTCTGGAAAAAGTTGGATTGGCTGGAAAAAGCTTCGATCTCTGATC 141
Db 11176 AAATACCGGGAGATTGAGAAAGTTGAAGAGTTGATGGAAATAACAAAGT 1235
Query 142 GTGGAAAGAGTTAACGTTAGCTGCTAACATGTTGATTGGAAAAGAGTTG 201
Db 1236 GATGAAATAATTTTATTATTGATGAAAGTACATCTTGATTGGCACGGCAGRA 1295
Query 202 GG 203
Db 1296 GG 1297

RESULT 15
US-08-961-527-115/C
; Sequence 115, Application US/08961527
; Publication No. US20042323A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:

US-08-961-527-115/C
Query 62 AAGGAAGATTCAATCCGAAAGGTGTGGAAAGTTGGCTGGAAAAAGTC 121
Db 105 AGACAGTAGTAAATTACTGAAAGATGTTGAGTTATCTGTTGAACTAATG 164

```
; CLASSIFICATION: 424
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 115:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11303 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-527-115

Query Match          10.4%; Score 32.4; DB 8; Length 11303;
Best Local Similarity      51.4%; Pred. No. 55;
Matches    75; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
Qy          62 AAAGAAGATTCATCATCCAGAAANGTGTGAAAGTGTGGATTGGCTGGAAAAGTC 121
Db          6373 AGACAAAGTTAAATTACTGAAAGATGGTTGAGTTATCNGTTGGAATCAAATGA 6314
Qy          1122 GATCTGACCATCTCTAGATCGTGGAAAAGTTACGTTACCTACATATGTGGA 181
Db          6313 GTTGAAACATCTTACTTTATGGAATTAACACGGCAGAAATTAAAGTCAGTTGTTT 6254
Qy          182 TTGTGAAAAAAGAGACGGAGGCATC 207
Db          6253 TTCTGAAAAAATTGAGTCGGATC 6228
```

Search completed: July 3, 2005, 07:33:13
Job time : 652.235 secs

Post-processing: Minimum Match 0\$ Maximum Match 100%									
Listing First 45 summaries									
Database :									
	EST:\$								
	1: gb_est1:\$								
	2: gb_est2:\$								
	3: gb_ntc:\$								
	4: gb_est3:\$								
	5: gb_est4:\$								
	6: gb_est5:\$								
	7: gb_est6:\$								
	8: gb_gss1:\$								
	9: gb_gss2:\$								
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.									
SUMMARIES									
Result No.	Score	Query Match	Length	DB ID	Description	FEATURES	Source	Location/Qualifiers	
1	38.8	12.4	777	BZ096570	BZ096570 CH230-236 CNS000IOV			1. .777	
c 2	37.8	12.2	1007	9	AL076752 Droophil				
c 3	37.8	12.1	1101	9	BZ000FBP				
c 4	37.2	11.9	754	8	BZ063129 11d06g03,				
c 5	36.6	11.7	925	9	AL074475 Droophil				
c 6	36.4	11.7	765	9	CCB07465 ZMMBc051				
c 7	36.4	11.7	2707	3	AK048004 Mus muscu				
c 8	36.2	11.6	633	7	CF979725 r98c09.y				
c 9	36	11.5	368	8	BZ840661 CH230-261				
c 10	36	11.5	576	9	CE352559 tiger-gss				
c 11	36	11.5	960	9	AL066226 Droophil				
c 12	36	11.5	1070	9	AG120841 Pan trogl				
c 13	35.8	11.5	367	8	BZ852956 CH240-217				
c 14	35.6	11.4	921	9	AG561995 Mus muscu				
c 15	35.6	11.4	1101	9	AL098595 Droophil				
c 16	35.4	11.3	507	9	CR498396 Medicago				
c 17	35.4	11.3	508	8	AQ439462 AQ5139862				
c 18	35.4	11.3	513	8	BX185081 Danio rer				
c 19	35.4	11.3	513	8	AZ570160 271cvB08				
c 20	35.4	11.3	585	9	BX198554 Danio rer				
c 21	35.4	11.3	1449	9	AG91030 Mus muscu				
c 22	35.2	11.3	375	9	CE224908 tiger-gss				
c 23	35.2	11.3	392	9	CE224907 tiger-gss				
c 24	35.2	11.3	567	4	BI667582 60329418				

CHORI-230 Rat (BN/SNHSc/MCW) BAC library produced by
Pieter de Jong"

Qy	95	AAGTTGTGGATTGCGTGGAAAAGTCGATCTGACATCTCTAGATCGTGAAAGAGTTC	154
Db	594	AAKTTTGTGKGKGKAATAWATKTGKGSTARGADARTNKNGAGAAKARAAAARGGG	535
Qy	155	ACGTTA	160
Db	534	AKCGEDA	529

RESULT	3	CNS00HBP	CNS00HBP	1101 bp	DNA	linear	GSS 03-JUN-1998
LOCUS		Drosophila melanogaster	Genome survey sequence	TET3	end of BAC:		
DEFINITION		BACR3A09 of RP1-98	library from Drosophila melanogaster (fruit fly)				
ACCESSION		AL073282	Genomic survey sequence.				
VERSION		AL073282.1					
KEYWORDS		GSS					
SOURCE		Drosophila melanogaster	(fruit fly)				
ORGANISM		Drosophila melanogaster					
EUKARYOTA; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydrioidae; Drosophilidae; Drosophila.							
REFERENCE		1	(bases 1 to 1101)				

RESULT	2	CNS00JOY	1007 bp	DNA	linear	GSS 03-JUN-1999
LOCUS		<i>Drosophila melanogaster</i>	genome survey sequence	ETR3	end of	BAC:
DEFINITION		BACR3B02 of RPCI-98 library from <i>Drosophila melanogaster</i> (fruit fly)	Genomic survey sequence.			
ACCESSION	AL076732	GI: 4996128				
VERSION						
KEYWORDS						
SOURCE		<i>Drosophila melanogaster</i> (fruit fly)				
ORGANISM		<i>Drosophila melanogaster</i>				
		Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Drosophyceza; Muscomorpha; Ephydriodea; Drosophilidae; Drosophila.				

REFERENCE : 1 (bases 1 to 1007)
 AUTHORS : Genoscope.
 TITLE : Direct Submission
 JOURNAL : Submitted (02-JUN-1999) Genoscope - Centre National de Sequencing
 COMMENT : (E-mail : seqref@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr
 determination of this BAC-level sequence was carried out as part of
 collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org>. The BDGP Drosophila melanogaster BAC library was prepared by Kazuyoshi Osoegawa and Aaron Mammone in Pieter Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial ECO RI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y²; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the PACBAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES	source	Location/Qualifiers	
		1. 1007	

Qy	15	GAAGGGCCCTGGTAATACCAATTCTGGAAAAAAGTTGGTAATACGGAAAAGAAAGATTCTAT 74
Db	741	GAARGGSSTTTGKRMWTTATCAATGYSASAGGGTTAKDAYCAAGGGMTTHTM 800
Qy	75	CATCGAGAAAAAGGTGTGAAAGTGTGTGAAAGTTGGTGGAAAAAGTCATCTGCACATCT 134
Db	801	TWTWTHAHHGATGTGCAWWADTTCTMNCARTCNGGAAAWAWATTKGARGTMTYTCATAAT 860
Qy	135	CTAGATCGGGAAAAAAGTCAAGCTTACGGCTTACCTACATATGTCATTGTTGGAAAAGA 194
Db	861	TWTCCGMXSGGGNMWTTGARAGCCCCGCTTYTTAWAWAAAARRAGMVTAWAMMAMA 920
Qy	195	AGACGGGCCATCGGGAAAAGAAGCTTGTACGGCTGAGCATAGTATAGTATAGTAC 254

AUTHORS	Bharti,A.K., Young,S., Kawchok,S., Keizer,G., Bronzino,A.C., Rouzard,K., Fuks,G., Yu,Y., Wing,R. and Messing,J.
TITLE	Sequencing of the maize genome at PGIR (2003b)
JOURNAL	Unpublished (2003)
COMMENT	Contact: Bharti,A.K. Dr.Joachim Messing's lab The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers University 190 Frelinghuysen Road, Piscataway, NJ 08854, USA Tel: 732 445 3801 Fax: 732 445 5755 Email: bharti@waksman.rutgers.edu Seq Primer: T7
FEATURES	High quality sequence start: 401. Location/Qualifiers 1. .765 /organism="Zea mays" /mol_type="genomic DNA" /cultivar="B73" /db_xref="taxon:15771" /clone="ZMMBBC0514M03" /lab_host="E. coli DH10B" /clone_lib="ZMMBBC" /note="vector: PTARBA1.3; Site_1: BamHI; Site_2: BamHI"
SOURCE	Query Match 11.7%; Score 36.4; DB 9; Length 765; Best Local Similarity 50.6%; Pred. No. 9.5; Matches 88; Conservancy 0; Mismatches 86; Indels 0; Gaps 0;
ORIGIN	Qy 81 GAAAAGGTCTGGAAAAAGGTGCGATTGGTGGAAAAAGGTCTGACCATCTCTAGAT 140 Db 491 GAAGAGAGCCGGAGATTGGGGATGAGAGCTAAACTAGCTAGAAGCTCT 550 Qy 141 GTGCAAALAGTTCACTGTTACGTGTTACATATGGATTGTGATTGTTCTGATTCAGT 200 Db 551 CTAAGATAAAATGGAAAGTAGAAGTTGATTGTTCTGATTCAGTCAAATGGCCGT 610 Qy 201 AGGATCGCTGGAAAAAGAGCTTGACCTGTAGCTGCTGAGTAGATAGATACT 254 Db 611 AGACCTCTTTATATAGGAGGGGGCTGACCCTTACCGACTGGATT 664
RESULT	AK048004 2707 bp mRNA linear HTC 03-APR-2004
LOCUS	Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone:c130029F03 product:unclassifiable, full insert sequence.
DEFINITION	
ACCESSION	AK048004
VERSION	GR26092584
KEYWORDS	HTC; CAP trapper.
SOURCE	Mus musculus (house mouse)
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	Carninci,P. and Hayashizaki,Y. 1. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 9927953 10349636
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499314
PUBMED	11042159
REFERENCE	3. Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
AUTHORS	Kondo,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishihine,T., Harada,A., Yamamoto,R., Matsunoto,H., Sakauchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsaura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	4. The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	5. The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6. (bases 1 to 2707)
AUTHORS	Adachi,J., Aizawa,K., Hanagaki,T., Hara,A., Hashizume,W., Fukuda,S., Furuno,M., Hiramoto,K., Hirooka,T., Hirozane,T., Hayashida,K., Hayatsu,N., Hisamatsu,K., Horii,F., Imotani,K., Ichinose,Y., Kondo,S., Konno,H., Kouda,M., Katchi,H., Kawai,J., Kojima,Y., Kondo,Y., Kourada,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numata,R., Ohno,M., Ohnato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sasaki,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sobage,Y., Tagami,M., Tagawa,A., Takaku,Akihira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE	Direct Submission (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-riego@riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
JOURNAL	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details.
FEATURES	URL: http://genome.gsc.riken.jp/ /location/Qualifiers 1. .2707 /organism="Mus musculus" /mol_type="mRNA" /strain="IC57BL/6J" /db_xref="FANTOM_DB_C130029F03" /db_xref="TAXON:10090" /clone="C130029F03" /tissue="head" /clone_id="RIKEN Full-length enriched mouse cDNA library" /dev_stage="16 days embryo" 1. .2707 /note="unclassifiable"
ORIGIN	Query Match 11.7%; Score 36.4; DB 3; Length 2707; Best Local Similarity 52.7%; Pred. No. 13; Matches 79; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

cycles were done using 1st strand and primers specific to SL1 leader sequence (SL1-NotI) and 3' end (XhoI-No-TR). dsDNA was digested using XhoI/NotI, fractionated on Chroma-spin 400 columns (Clontech) and ligated to digested pGEM -1zf(+) plasmid. Chemically competent cells were used as host. This library was provided by David Bird's lab (david_bird@ncsu.edu), of North Carolina State University."

RESULT 8

Qy 209 GTGGAAAGAAGCTGTACCGTGTACGTTGACGATAGATACACCGTCACGGTCC 268
 LOCUS CP979725 633 bp mRNA linear EST 25-NOV-2003
 DEFINITION r980c05.Y1 Meloidogyne hapla female SL1 pgem Meloidogyne hapla cDNA 2264
 5' similar to TR_Q9XW8 Y37DPA.14 PROTOIN. [1] ; mRNA sequence.
 CF979725

ACCESSION Qy 11 ARRAGAAAAGGCCCTGGTAATACCATGGAAAAAGTTGGTAAATACGGAAAAAGAAGAT 70
 VERSION 92; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
 KEYWORDS Db 449 ATGCACAAATTGCTGGATGAACTTTCGATGAAGTTGAGGAAATAATAGAA 508
 SOURCE Qy 71 TCATCATCCAGAAGAGGTGTGGAAAGTGTGGATTGCTGGAAAGTTGATCTGACC 130
 ORGANISM Meloidogyne hapla Db 509 AGAAGTGGTAATTTTGTAGGGTATTGTAATAATGATTAAATTAAATTTC 568
 Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
 Tylenchoidea; Heteroderidae; Meloidogyinae; Meloidogyne.
 1 (bases 1 to 633)

REFERENCE Cf979725..1 GI:38512774

AUTHORS McCarter,J., Clifton,S., Chiappelli,B., Pape,D., Martin,J.,
 Bowers,Y., Dantes,M., Marra,N., Hillier,L., Kucaba,T., Theising,B.,
 Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
 Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
 Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
 McCann,R., Waterston,R. and Wilson,R...

TITLE Unpublished (1999)

JOURNAL Contact: McCarter,JP

COMMENT The Washington Univ. Nematode EST Project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

Cloned unidirectionally. Poly (A)+ RNA was concentrated and purified using Dynabeads (Dynal) and mRNA eluted for 1ST strand synthesis. 1st strand cDNA created using MMLV RT (PowerScript, Clontech) and primed with Oligo(dT) with XhoI site (primer CD5II/(3'-XhoI)) and 5', SMART anchor added using chimeric DNA-RNA oligo (SMART-NotI-r-GGG). SMART-NotI-r-GGG was used in case a PCR using the SMART template was also desired. 15 PCR cycles were done using 1st strand and primers specific to SL1 leader sequence (SL1-NotI) and 3' end (XhoI-No-TR). dsDNA was digested using XhoI/NotI, fractionated on Chroma-spin 400 columns (Clontech) and ligated to digested pGEM -1zf(+) plasmid. Chemically competent cells were used as host. This library was provided by David Bird's lab (david_bird@ncsu.edu), of North Carolina State University.

Seq primer: Sp6.

FEATURES source

1. .633
 /organism="Meloidogyne hapla"
 /mol_type="mRNA"
 /db_xref="taxon:6305"
 /sex="female"
 /tissue_type="whole organism"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Meloidogyne hapla Female SL1 PGEM"
 /note="Vector: pGEM-1zf(+)(Promega); Site_1: XhoI;
 Site_2: NotI; Cloned unidirectionally. Poly (A)+ RNA was concentrated and purified using Dynabeads (Dynal) and mRNA eluted for 1ST strand synthesis. 1st strand cDNA created using MMLV RT (PowerScript, Clontech) and primed with Oligo(dT) with XhoI site (primer CD5II/(3'-XhoI)) and 5', SMART-NotI-r-GGG. SMART-NotI-r-GGG was used in case a PCR using the SMART template was also desired. 15 PCR

cycles were done using 1st strand and primers specific to SL1 leader sequence (SL1-NotI) and 3' end (XhoI-No-TR). dsDNA was digested using XhoI/NotI, fractionated on Chroma-spin 400 columns (Clontech) and ligated to digested pGEM -1zf(+) plasmid. Chemically competent cells were used as host. This library was provided by David Bird's lab (david_bird@ncsu.edu), of North Carolina State University."

ORIGIN

Query Match 11.6%; Score 36.2; DB 7; Length 633;
 Best Local Similarity 49.7%; Prod. No. 10;
 Matches 92; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

Qy 11 ARRAGAAAAGGCCCTGGTAATACCATGGAAAAAGTTGGTAAATACGGAAAAAGAAGAT 70
 Db 449 ATGCACAAATTGCTGGATGAACTTTCGATGAAGTTGAGGAAATAATAGAA 508
 Qy 71 TCATCATCCAGAAGAGGTGTGGAAAGTGTGGATTGCTGGAAAGTTGATCTGACC 130
 Db 509 AGAAGTGGTAATTTTGTAGGGTATTGTAATAATGATTAAATTAAATTTC 568
 Db 510 AGAAGTGGTAATTTTGTAGGGTATTGTAATAATGATTAAATTAAATTTC 568
 Qy 131 ATCTAGATGTCGTGAAAGTCACTTACCTACATATGTCGATTGCGAA 190
 Db 569 CTCTTAAAGTAANAAATAATTAGAACAAATTGTTGATTAANTTTTTTAAAA 628
 Qy 191 AAGGA 195
 Db 629 AAAA 633

RESULT 9

B2840661 LOCUS B2840661_261B20.TJ CHORI-240 Bos taurus genomic clone CH240_261B20_368 bp DNA linear GSS 18-MAR-2003

DEFINITION Genomic survey sequence.

ACCESSION B2840661

VERSION B2840661..1 GI:29068020

KEYWORDS GSS,
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Bovidae;
 Bovine; Bos.

REFERENCE 1 (bases 1 to 368)

AUTHORS Zhao,S., Shetty,J., Sharman,S., Tsugaya,G., Geer,K., Shwartsbeyn,A., Gebregeorgis,E., Chen,D., Riggs,F., de Jong,P., Crawford,A.M. and McEwan,J.C.

TITLE Bovine BAC End Sequences from Library CHORI-240

JOURNAL Unpublished (2003)

COMMENT Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org

Clones are derived from the bovine BAC library CHORI-240 (<http://www.chori.org/bapac/bovine240.htm>). For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (<http://www.chori.org/bapac/ordering-information.html>). This work was undertaken as part of the International Bovine BAC Mapping Consortium (IBMC) by AgResearch Ltd., New Zealand and The Institute of Genomic Research (TIGR), USA.

Plate: 261 row: B column: 20
 Seq primer: SP6
 Class: BAC ends
 Location/Qualifiers FEATURES source

1..368
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="breed: Hereford"

/db_xref="taxon:9913"
 /clone="CH240_261B20"
 /sex="Male"
 /cell_type="Blood"
 /clone_lib="CHORI-240"
 /note="Vector: pTARBAC1.3; Site_1: MboI; Site_2: NheI; Hereford bull Li Domino 99375; CHORI-240 Bovine BAC library (Male) produced by Pieter de Jong"

ORIGIN

Query Match	Score 36;	DB 8;	Length 368;
Best Local Similarity	49.0%	Pred. No. 10;	
Matches	96;	Mismatches	100;
Conservative	0;	Indels	0;
Gaps	0;		

Qy 35 ATTGTGAAAGTGGTATACCGAAAAGATTCACTCAGAAAAGGTGTGAA 94
 Db 143 ACTGAAAAAATAGATAATAAATAAAAAAACAAAAGAAATGGAGTTT 202
 Qy 95 AAGTGTGGATTGGCTGGAAAAGTCGATCTGACCATCTCTAGATCCGGAAAAGTC 154
 Db 203 CTCACGTATTCCTGTAGATCTGGTTCTCTGTGAGCAGGAGTAGT 262
 Qy 155 ACGTAGCGCTTACTACATATGTGGATGTGGAAAAGAACGGGACCATGGTCAA 214
 Db 263 GAGTACCTAAACGCAAGACTGTCTCATGTCAGAGGCTTAATTCTGTTTC 322
 Qy 215 AAAGAGCTGTACG 230
 Db 323 AAAGATCCCCGTGCA 338

RESULT 10
 CE352559/C CB352559 576 bp DNA linear GSS 26-SEP-2003
 DEFINITION tigr-dbs-dog-17000361425765 Dog Library Canis familiaris genomic, genomic survey sequence.

ACCESSION	CE352559
VERSION	1
KEYWORDS	Gi:36188764
SOURCE	Canis familiaris (dog)
ORGANISM	Canis familiaris

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Pissipedia; Canidae; Canis.
 REFERENCE 1 (bases 1 to 576)
 AUTHORS Kirckness,E.F., Bafna,V., Halpern,A.I., Levy,S., Remington,K., Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and Venter,J.C.
 TITLE The dog genome: survey sequencing and comparative analysis
 JOURNAL Science 301 (5641), 1898-1903 (2003)
 JOURNAL 14512627
 PUBMED 22875432
 COMMENT Contact: Kirkness EF
 The Institute for Genomic Research
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
 Rockville, MD 20850, USA
 Tel: 301-838-0200
 Fax: 301-838-0208
 Email: ekirkness@tigr.org

FEATURES SOURCE

Class:	shotgun.
Location/Qualifiers	1..576
/organism="Canis familiaris"	
/mol_type="genomic DNA"	
/strain="Standard Poodle"	
/db_xref="taxon:9615"	
/clone_lib="Dog Library"	
/note="Site 1: BstXI; Libraries were prepared from peripheral blood"	

ORIGIN

Query Match	Score 36;	DB 9;	Length 576;
Best Local Similarity	54.5%	Pred. No. 12;	
Matches	72;	Mismatches	60;
Conservative	0;	Indels	0;
Gaps	0;		

Qy 20 GGCTCTGGTAATAACCATATGGAAAAGTGGTAAATACGGAAAAAGAGATTCATCATCC 79
 Db 210 GGAGTGGGACATAAGCGAGAAATATTGGGTGAGGGATAGTATTGCCAAATA 151
 Qy 80 AGAAAGCTGGAAAGTTGGATGCTGGAAAGTGTGACCATCTCTAGA 139
 Db 150 AGAAAGCTGGAAACTAGAGATGAGTGGAACTAGAAATGAGTGGTCCAA 91
 Qy 140 TCGTGGAAAAG 151
 Db 90 GCATAGGATAAG 99

RESULT 11
 CNS0062M/C CNS0062M LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC # BAC14J16 of RPCI-38 library from Drosophila melanogaster (fruit fly); genomic survey sequence.

ACCESSION	AL066226																																				
VERSION	AL066226..1																																				
KEYWORDS	GI:4945939																																				
ORGANISM	Drosophila melanogaster																																				
COMMENT	Drosophila melanogaster (fruit fly) GSS. Drosophila melanogaster Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydriidae; Drosophilidae; Drosophila. REFERENCE 1 (bases 1 to 960) AUTHORS Genoscope. TITLE Direct Submission JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequenage : BP 191 91006 EVRY Cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr.) DEFINITION Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org. The BDGP Drosophila melanogaster BAC library was prepared by Kazuyoshi Osoawa and Aaron Mammober in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain v2; cn bw sp, the same strain used for the BDGP P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm. <p>FEATURES SOURCE</p> <table border="0"> <tr> <td>organism="Drosophila melanogaster"</td> <td></td> </tr> <tr> <td>/mol_type="genomic DNA"</td> <td></td> </tr> <tr> <td>/db_xref="taxon:7227"</td> <td></td> </tr> <tr> <td>/clone="BACR14J16"</td> <td></td> </tr> <tr> <td>/clone_id="RPCI-98"</td> <td></td> </tr> <tr> <td>/notes="end : T7"</td> <td></td> </tr> </table> <p>ORIGIN</p> <table border="0"> <tr> <td>Query Match</td> <td>Score 36;</td> <td>DB 9;</td> <td>Length 960;</td> </tr> <tr> <td>Best Local Similarity</td> <td>28.1%</td> <td>Pred. No. 13;</td> <td></td> </tr> <tr> <td>Matches</td> <td>38;</td> <td>Conservative</td> <td>49;</td> </tr> <tr> <td></td> <td></td> <td>Mismatches</td> <td>48;</td> </tr> <tr> <td></td> <td></td> <td>Indels</td> <td>0;</td> </tr> <tr> <td></td> <td></td> <td>Gaps</td> <td>0;</td> </tr> </table> <p>Qy 91 GGAAGCTGGATGCTGGAAAAGTGTGATCTGACCATCTCTAGATCTGGAAAAA 150 Db 913 GAGGMMGGGMMGGGCTKGCHHNMGMGHTGGMSTMVGMGMMTMGGMTGGTNGT 854 Qy 151 GPTCACGTAGCTAGCTAGTACATGTGGATTTGCAAAAGAGAGCAGGGCATGGT 210 Db 853 MRTTGMGTTMGGMGMMGTTTGMGRTVGMGMMKMMKMMTGGGG 794 Qy 211 CGAAAAAGAAGCTTG 225</p>	organism="Drosophila melanogaster"		/mol_type="genomic DNA"		/db_xref="taxon:7227"		/clone="BACR14J16"		/clone_id="RPCI-98"		/notes="end : T7"		Query Match	Score 36;	DB 9;	Length 960;	Best Local Similarity	28.1%	Pred. No. 13;		Matches	38;	Conservative	49;			Mismatches	48;			Indels	0;			Gaps	0;
organism="Drosophila melanogaster"																																					
/mol_type="genomic DNA"																																					
/db_xref="taxon:7227"																																					
/clone="BACR14J16"																																					
/clone_id="RPCI-98"																																					
/notes="end : T7"																																					
Query Match	Score 36;	DB 9;	Length 960;																																		
Best Local Similarity	28.1%	Pred. No. 13;																																			
Matches	38;	Conservative	49;																																		
		Mismatches	48;																																		
		Indels	0;																																		
		Gaps	0;																																		

DB	793	GGNNMMGGMVGMGMG	779	
genomic survey sequence.				
RESULT	1.2			
LOCUS	AG120841/c	AG120841	1070 bp	DNA linear GSS 04-NOV-2001
DEFINITION	Pan troglodytes DNA, clone: PTB-129L09.F, genomic survey sequence.			
VERSION	AG120841	AG120841.1	GI:16650006	
KEYWORDS	GSS.			
ORGANISM	Pan troglodytes (chimpanzee)			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.				
REFERENCE	1			
AUTHORS	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.			
TITLE	BAC end sequences of Library PTB			
JOURNAL	Unpublished			
REFERENCE	2	(bases 1 to 1070)		
AUTHORS	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.			
TITLE	Direct Submission			
JOURNAL	Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Sheihiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan. (E-mail: chimpdb@gsr.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: +81-45-503-9111, Fax: +81-45-503-9170)			
COMMENT	Clones are derived from the chimpanzee BAC library PTB. This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.			
PRIMERS	Sequencing: -21M13			
LIBRARY	Vector : pKS145			
	R.Size 1 : SacI			
	R.Size 2 : SacI.			
FEATURES	source	1..1070	Location/Qualifiers	
			/organism="Pan troglodytes"	
			/mol_type="genomic DNA"	
			/db_xref="taxon:913"	
			/clone="CH240_217K24"	
			/sex="Male"	
			/cell_type="Blood"	
			/clone_lib="CHORI-240"	
			/note="Vector: pTARBAC1.3; Site_1: MboI; Site_2: MboI; Hereford bull Li Domino 99375; CHORI-240 Bovine BAC library (Male) produced by Pieter de Jong"	
ORIGIN				
Query	44	AAAGTGTGTTAATACGGAAAAAGAGATTCACTCCAGAAAAGGTGTGAAAGTTCTGG 103	Match Score 11.5%; Best Local Similarity 48.5%; Matches 99;	Length 1070; Pred. No. 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	1070	AAATAAGGGAAAGGGAAAGGGGAAATAGGAAAGGATAGATGGA 1011		Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	104	ATTCAGCTGAAAGAGTTGATCTGACCTCTAGATCTGTGAAAGTTCTCGTTAGCG 163		
Db	1010	ATAGGGAGAAAAAAAGGAGATAAGATAAAAGGAGGAGGGGGGGGT 951		
Qy	164	CTTACGTPACATATGTGATTGTGAAAGAACGAGGCTAGAGCT 223		
Db	950	CGAAAAGGGAAGGGTGGAGGAAAGGAAGAAGGAAGGA 891		
Qy	224	TGTACGCTGTACTCTGAGATAGA 247		
Db	890	GGGGGGGACAGAAAGAGAGAGA 867		
RESULT	14	AG561955	LOCUS	AG561955
DEFINITION	Mus musculus molossinus DNA sequence.		DEFINITION	Mus musculus molossinus DNA sequence.
ACCESSION	AG561955		ACCESSION	AG561955
VERSION	1		VERSION	1
KEYWORDS	GSS.		KEYWORDS	GSS.
RESULT	1.3	AG561956	LOCUS	AG561956
DEFINITION	CH240_217K24.TV CHORI-240 Bos taurus genomic clone CH240_217K24.		DEFINITION	CH240_217K24.TV CHORI-240 Bos taurus genomic clone CH240_217K24.
ACCESSION	MSNG01-481K04.r7		ACCESSION	MSNG01-481K04.r7
VERSION	1		VERSION	1
KEYWORDS	GSS.		KEYWORDS	GSS.

SOURCE	Mus musculus molossinus	AUTHORS	Genoscope
ORGANISM	Mus musculus molossinus	TITLE	Direct Submission
REFERENCE	Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;	JOURNAL	Centre National de Sequencage :
1	Eutheria; Sciuromorpha; Muridae; Murinae; Mus.		BP 191 91.006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)
AUTHORS	Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.	COMMENT	- Web : www.genoscope.cns.fr
TITLE	Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.		Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (BDGP) -
JOURNAL	Unpublished		http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Geneviève Payan. It has been constructed in the vector pBeloBAC11.
REFERENCE	2 (bases 1 to 921)	FEATURES	Location/Qualifiers
AUTHORS	Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.	Source	1..1101
TITLE	Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC);		/organism="Drosophila melanogaster"
JOURNAL	1-7-22 Suehiro-Chou Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@riken.jp, URL: http://hgp.gsc.riken.go.jp/ , Tel:81-45-503-9111, Fax:81-45-503-9170)		/mol_type="genomic DNA"
COMMENT	Copies are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kunuya Abe (abe@rtc.riken.jp).		/db_xref="taxon:7227"
	Tsukuba Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN) 3-1-1, Koyadai, Tsukuba, 305-0074 Japan phone: 81-298-36-9189, fax: 81-298-36-9199 e-mail: abe@rtc.riken.jp		/clone="BACN0K20"
PRIMERS			/clone_id="Bac8BAC"
			/plasmid="pBeloBAC11"
			/note="end : T7"
SEQUENCING	: T7	ORIGIN	
LIBRARY		Query Match	Score 35.6; DB 9; Length 1101;
Vector	: pBAC3.6	Best Local Similarity	11.4%; Pred. No. 18;
R.Site 1	: EcoRI	Matches	34; Mismatches 76; Indels 0; Gaps 0;
R.Site 2	: EcorI.	Qy	36 TTCTGGAAAAAAACTTGGTAATACGGAAAAGGAGATTCTCATCCAGAAAAGGTGTGGAAA 95
FEATURES	Location/Qualifiers	Db	818 KRKKKDRAWDKRRKDKDAAKAKADAAAGDKAKRRAAGDKAKAKAKAKAKDD 877
Source	1..921	Qy	96 AGTTGTGAAATTGCGTGGAAAAGTCGATCNSACACATCTAGATCGTGGAAAAGCTCA 155
	/organism="Mus musculus molossinus"	Db	878 DDAAATAKKATKAEDKAKKCKKKDKDAKAKAKADAAKDDDKDDKDDKDDAKA 937
	/mol_type="genomic DNA"	Qy	156 CGTAGGCCTTAGTGTACATATGTGATTGTGGAAAAGAACGGGAGCATGGTGGAAA 215
	/sub_species="molossinus"	Db	938 DRKKWKRDRADKAKKDDDKDKAKDKDDKDDKDDKDDKDDKDDKDDKDDKDDKDD 997
	/db_xref="taxon:57486"	Qy	216 AAGAAGCTTGACGCTGTACGTGACGGATAGATAGA 251
	/clone="MSMg01-481K04.T7"	Db	998 DDAADAKAKADAKAKAKADAKAKAKADAKAKAKADAKAKAKADAKAKAKADAKA 1033
	/sex="male"	Search completed: July 3, 2005, 04:51:39	
	/tissue type="mixture of kidney and spleen"	Job time : 3933.41 secs	
	/clone_id="Bac8BAC"		
SEQUENCING	: T7		
LIBRARY			
Vector	: pBAC3.6		
R.Site 1	: EcoRI		
R.Site 2	: EcorI.		
FEATURES	Location/Qualifiers		
Source	1..921		
	/organism="Mus musculus molossinus"		
	/mol_type="genomic DNA"		
	/sub_species="molossinus"		
	/db_xref="taxon:57486"		
	/clone="MSMg01-481K04.T7"		
	/sex="male"		
	/tissue type="mixture of kidney and spleen"		
	/clone_id="Bac8BAC"		
SEQUENCING	: T7	ORIGIN	
LIBRARY		Query Match	Score 35.6; DB 9; Length 921;
Vector	: pBAC3.6	Best Local Similarity	43.4%; Pred. No. 17;
R.Site 1	: EcoRI.	Matches	83; Mismatches 0; Indels 0; Gaps 0;
R.Site 2	: EcorI.	Qy	53 ATATCGGAAAGAGAGATCATCTCCAGAAACCTGGGAAGAGTTGTGGATTTGGCTGG 112
FEATURES	Location/Qualifiers	Db	472 AAAAAGAAAAAGGGGNAAAAGGGGAAAAGGGGGAAAANCGGGGGGGGGGGGAG 531
Source	1..921	Qy	113 AAAAGTTCGATCTGACCATCTAGATCGGGAAAAGACTCAAGTTACGTGCTACGTAC 172
	/organism="Mus musculus molossinus"	Db	532 AAAAAGAAAAAGGGGAAAGGGGAAAGGGGAAAGGGGAAAGGGGAAAGGGGAG 591
	/mol_type="genomic DNA"	Qy	173 ATATGTGGATTGTGGAAAAAGAACGAGGGCATCGTGGAAAAGACTCAAGTTACGTAC 220
	/sub_species="molossinus"	Db	592 AGGGGGGGGGGGAAAATAATAAGGGGGGGGGAAAAGGGGGGGGGGGGGGGGGGGGG 639
SEQUENCING	: T7	RESULT 15	
LIBRARY		CNS0106X	CNS0106X
Vector		DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC
R.Site 1			BACN0K20 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
R.Site 2		ACCESSION	AL098535.1 GI:5610205
		VERSION	AL098535
		KEYWORDS	Drosophila melanogaster (fruit fly)
		ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscicomorpha; Ephydioidea; Drosophilidae; Drosophila; Drosophila
		REFERENCE	1 (bases 1 to 1101)

Result No.	Score	Query Match	Length	DB ID	Description
1	130	100.0	452	3 AAA28449	Aaa28449 Artificial
2	89	68.5	89	3 AAA28445	Aaa28445 Synthetic
3	63.6	48.9	3212	5 AAP30968	Aaa30968 Aspergillus
4	63.6	48.9	3288	5 AAF30970	Aaa30970 Aspergillus
5	63.6	48.9	3329	6 AAF30969	Aaa30969 Aspergillus
6	61	46.9	1006	2 AAX81436	Aaa81436 Arabidopsis
7	60.4	46.5	1138	2 AAT85655	Aaa85655 pMj01 DNA
8	60.4	46.5	1138	2 AAT59713	Aaa59713 Plasmid P
9	60.4	46.5	1138	2 AAX7885	Aax7885 Plasmid P
10	59.6	45.8	157	6 AAD44423	Aad44423 pGL482 co
11	58.2	44.8	2200	6 ABLS5882	Abi5882 HPV16-L1
12	56.2	43.2	5033	2 AAQ7465	Aaq7465 pHL1406
13	52	40.0	6395	2 AAQ95155	Aaq95155 Tobacco m
14	52	40.0	6395	2 AAZ20642	Aaz20642 TMV-based
15	52	40.0	6395	3 AAC62369	Aac62369 cDNA sequ
16	52	40.0	6395	4 AAF82330	Aaf82330 Wild-type
17	52	40.0	6425	2 AAZ20644	Aaz20644 TMV-based
18	52	40.0	6425	4 AAF82332	Aaf82332 Tobacco m
19	52	40.0	6439	2 AAZ20643	Aaz20643 TMV-based
20	52	40.0	6439	4 AAF82331	Aaf82331 Tobacco m

PF	25-FEB-1999;	99EP-00301419.
XX	09-NOV-1998;	98IN-DE003322.
PR		(COUNL) CSIR COUNCIL SCI IND RES.
XX		
PA		
XX		
PI	Tuli R,	Sawant SV,
XX		Singh PK,
DR		Gupta SK;
XX		
WPI:	2000-341712/30.	
PT	New chemically synthesized artificial promoter, useful high level expression of transgenes in different organisms.	
XX		
PT	New chemically synthesized artificial promoter, useful high level expression of transgenes in different organisms.	
XX		
PS	Claim 13; Page 11; 40pp; English.	
PS	Chemically synthesized artificial promoters are new and comprise a DNA sequence designed for a targeted level and pattern of gene expression by strategically putting together several signature sequences identified by sequence alignment and statistical analysis of a large database constructed for this purpose. A method for chemically synthesizing an artificial promoter for expressing genes at a desired level in different organisms is also claimed. The high level expression in a plant using such an artificial promoter (e.g. AAA28419) can be measured comprising polyethylene glycol (PEG) mediated transformation of plant protoplasts as well as biolistic mediated transformation of plant tissues including root, stem, intact leaf tissue followed by transient GUS assay to compare with a natural CaMV 35S promoter showing the desired level of activity. The promoter is useful for high level expression of transgenes in different organisms and for testing high level gene expression in plants (claimed). The promoter is biologically active and is efficient and can be synthesized to express in even the most complex organisms	
XX		
SQ	Sequence 452 BP; 157 A; 88 C; 95 G; 112 T; 0 U; 0 Other;	
Query Match	Score 130; DB 3; Length 452;	
Best Local Similarity	100.0%;	
Matches	130;	
Conservative	0;	
Mismatches	0;	
Indels	0;	
Gaps	0;	
Qy	1 TCACTATATAGGAAGTCATTCAATTGATGGCACCGTGTGATTTCTCACAA	60
Db	313 TCACTATATAGGAAGTCATTCAATTGATGGCACCGTGTGATTTCTCACAA	372
Qy	61 TTACCAACAAACAAACATATACAAATTATTAATCATATCTAG	120
Db	373 TTACCAACAAACAAACATATACAAATTATTAATCATATCTAG	432
Qy	121 ATAAACATG 130	
Db	433 ATAAACATG 442	
Query Match	Score 130; DB 3; Length 89;	
Best Local Similarity	100.0%;	
Matches	89;	
Conservative	0;	
Mismatches	0;	
Indels	0;	
Gaps	0;	
Qy	39 ACGTGTTCTATTCTCAAAATTACACAAACAAACAAACATATACAAAT	98
Db	1 ACGTGTTCTATTCTCAAAATTACACAAACAAACAAACATATACAAAT	60
Qy	99 TACTATTACAATTACATCTAGATAAACAA 127	
Db	61 TACTATTACAATTACATCTAGATAAACAA 89	
RESULT	3	
ID	AAF30968	standard; cDNA; 3212 BP.
XX		
AC	AAF30968;	
XX		
AC	AAF30968;	
XX		
DT	23-JUL-2001	(first entry)
XX		
DE	Aspergillus niger beta-glucosidase BGL1 cDNA in plasmid pJDB1.	
XX		
KW	Beta-glucosidase; BGL1; aroma; alcoholic beverage; ethanol;	
KW	transgenic plant; tobacco; pJDB1; ss.	
XX		
OS	Aspergillus niger.	
FH	Key	Location/Qualifiers
FT	mat_peptide	391 .2916
XX		
PN	/*tag=	a
XX		
PR	WO200136586-A2.	
XX		
PD	24-MAY-2000.	
XX		
PP	25-MAY-2001.	
XX		
PR	15-NOV-2000; 99EP-00301419.	
XX		
PR	09-NOV-1998; 98IN-DE003322.	
XX		
EP1002869-A1.		
XX		
PN		
XX		
AC	AAA28445;	
XX		
DT	29-AUG-2000	(first entry)
XX		
XX	Synthetic promoter 5' untranslated leader sequence.	
KW	Artificial promoter; 5' untranslated leader sequence;	
KW	transgene expression; plant; ss.	
XX		
OS	Synthetic.	
XX		

PR 19-NOV-1999; 9UUS-0044338.
 XX
 PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX
 PI Shu W, Marton I, Siegel DL, Ben-Ami B, Dekel M, Shoeyov O;
 XX
 DR WPI: 2001-355624/37.
 DR P-PSDB; AAB82326.
 XX
 PR New nucleic acid encoding beta-galactosidase from Aspergillus niger, useful e.g. for increasing aroma compound formation during alcoholic fermentation.
 PR Disclosure; Page 94-96; 106pp; English.
 CC The present sequence is that of Aspergillus niger Bl (CMI CC 324626) cDNA in plasmid pJDBl encoding the mature portion (see AAB8326) of beta-glucosidase BGl1. The plasmid was constructed by isolating the bg1 cDNA from Plasmid pETB1 using NcoI and BamHI and inserting it into pJD330 between the 35S promoter sigma fragment and the nos terminator, eliminating the gus gene. The gene cassette in pJDBl was isolated and inserted into binary vector pBINplus. Disarmed Agrobacterium tumefaciens LBA404 was transformed with the resulting vector and used to transform tobacco cv. Samson plants. The gene expression cassette provided for cytoplasmic expression of mature A. niger BGl1 in tobacco plants. BGl1 polymucleotides, polypeptides, host cells and transgenic plants are useful for fermentation, particularly to increase the level of aroma compounds in alcoholic drinks, as well as other fermentation products of plant material, hydrolyzing cellulose and thus increasing the level of fermentable glucose, increasing the production of alcohol, such as ethanol, from plant material, increasing the aroma released from a plant or a plant product, and increasing the hydrolysis of glycosides. Unlike previously known beta-glucosidases, BGl1 is very stable over a wide range of temperatures and pH, e.g. 68% retention of activity after 1 hour at 60 degree C
 XX Sequence 3212 BP; 771 A; 845 C; 858 G; 738 T; 0 U; 0 Other;
 SQ Query Match 48.9%; Score 63.6; DB 5; Length 3212;
 Best Local Similarity 87.7%; Pred. No. 6.2e-06;
 Matches 93; Conservative 0; Mismatches 9; Indels 4; Gaps 2;
 Qy 12 AGGAGTTCATTCTCATTTGGAAATGACA--CGTGTTGTCAATTCTCAAAATTACACA 69
 Db 274 AGGAGTTCATTCTCATTTGGAAATGACAAGCTTCTGAGATCCCTCAAAATTACACA 333
 Qy 70 ACAACAAACAAACAAACATTACAAATTACTATTACAAATTACA 115
 Db 334 ACAACAAACAAACAAACACAT--TACAAATTACAAATTACA 377

RESULT 4
 AAF30970 ID AAF30970 standard; cDNA; 3288 BP.
 XX
 AC AAF30970;
 XX DT 23-JUL-2001 (first entry)
 DE Aspergillus niger beta-glucosidase BGl1 cDNA in plasmid pJDCB1T.
 KW Beta-glucosidase; BGl1; aroma; alcoholic beverage; ethanol;
 KW transgenic plant; tobacco; pJDCB1T; Cell; signal peptide;
 KW endoplasmic reticulum; ss.
 OS Aspergillus niger.
 OS Unidentified.
 OS Synthetic.
 OS Chimeric.
 XX EH Key 391 . 3042
 FT Location/Qualifiers /*tag= a
 FT

KW Beta-glucosidase; BGL1; aroma; alcoholic beverage; ethanol;
transgenic plant; tobacco; pJDDB1; 8B.

OS Aspergillus niger.
Unidentified.
Chimeric.

XX
Key
PH Location/Qualifiers
CDS 391..3033
/tag= a
FT sig_peptide 391..507
/tag= b
FT mat_peptide 508..3030
/tag= c
FT /note= "BGL1 mature polypeptide"
PN WO200136586-A2.
XX
PD 25-MAY-2001.
XX
PF 15-NOV-2000; 2000WO-IL000758.
XX
PR 19-NOV-1999; 99US-00443338.
XX
PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
PX Shu W, Marton I, Siegel DR, Ben-Ami B, Dekel M, Shoebyov O;
PT DR; 2001-355624/37.
XX
DR P-BSDB; AAB2327.
XX
PR New nucleic acid encoding beta-galactosidase from Aspergillus niger,
useful e.g. for increasing aroma compound formation during alcoholic
fermentation.
XX
PS Disclosure; Page 97-99; 106pp; English.

CC The present sequence is that of a cDNA construct in plasmid pJDDB1
encoding a polypeptide (see AAB8327) comprising the C-terminal signal peptide
and the mature region of Aspergillus niger B1 (CMI CC 324626) beta-
glucosidase BGL1. The cDNA is placed under the control of the 3SS
promoter and nos terminator. The gene cassette in pJDDB1 was inserted
into binary vector pBNplus. Disarmed Agrobacterium tumefaciens LB4404
was transformed with the resulting vector and used to transform tobacco
cv. Samson plants. The gene expression cassette provided for secretion of
BGL1 into the apoplast. BGL1 polymeric peptides, host cells
and transgenic plants are useful for fermentation, particularly to
increase the level of aroma compounds in alcoholic drinks, as well as
other fermentation products of plant material, hydrolysing cellulose and
thus increasing the level of fermentable glucose, increasing the
production of alcohol, such as ethanol, from plant material, increasing
the aroma released from a plant or a plant product, and increasing the
hydrolysis or transglycosylation of glycosides. Unlike previously known
beta-glucosidases, BGL1 is very stable over a wide range of temperatures
and pH, e.g. 68% retention of activity after 1 hour at 60 degree C
SQ Sequence 3329 BP; 790 A; 882 C; 883 G; 774 T; 0 U; 0 Other;

Query Match 48.9%; Score 63.6; DB 5; Length 3329;
Best Local Similarity 87.7%; Pred. No. 6.2e-06;
Matches 93; Conservative 0; Mismatches 9; Indels 4; Gaps 2;

RESULT 7
AAT85665 standard; DNA; 1138 BP.
ID AAT85665
XX
AC
DT 25-AUG-1999 (first entry)
XX Arabidopsis thaliana promoter fragment produced by PCR.
DE Plant promoter; minimal promoter; transcription activating element;
KW KW
KW gene expression; transgenic plant; 8B.
XX Synthetic.
OS Arabidopsis thaliana.
XX OS
PN WO931258-A1.
XX PN
PD 24-JUN-1999.
XX PF 10-DEC-1998; 98WO-BP008162.
XX PR 12-DEC-1997; 97EP-00203912.
XX PA (MOGE-) MOGEN INT NV.
XX PI Stuiver MH, Sijbolds EH;
XX DR WPI; 1999-395188/33.
XX PT Novel constitutive plant promoters.
XX PS Example 10; Page 36-37; 44pp; English.
XX The specification describes plant promoters, where elements from one
CC promoter which are responsible for specific expression are combined with
CC elements from a second promoter which are responsible for a complementary
CC expression pattern. The novel promoter then shows expression in the
CC tissues and developmental stages which form the expression pattern of
CC both promoters. The plant promoter is characterized in that it comprises
CC a minimal promoter and transcription activating elements from a set of
CC promoters, the elements having a complementary pattern and level of
CC transcription in plant. The promoters of the invention are used to
CC express genes in transgenic plants. The present sequence was used in the
course of the invention
SQ Sequence 1006 BP; 320 A; 182 C; 183 G; 321 T; 0 U; 0 Other;
Query Match 46.9%; Score 61; DB 2; Length 1006;
Best Local Similarity 88.3%; Pred. No. 2.4e-05;
Matches 91; Conservative 0; Mismatches 5; Indels 7; Gaps 2;
XX
Oy 12 AGGAGTCAATTCAATTGGATGGAGCGGTGTTGATTTCTAACATTACAAAC 71
Db 906 AGGAGTCAATTCAATTGGAGGAGCACGAT---TTTACAAACATTACAAAC 960
XX
Oy 72 AACAAACAAACAAACAAACATTACATTACATTAACTTACATTAC 114
Db 961 AACAAACAAACAAACAT--TACAAATTACATTAACTTACATTAC 1001
XX
RESULT 8
AAT85665 standard; DNA; 1138 BP.
ID AAT85665
XX
AC
DT 21-NOV-1997 (first entry)
XX PMJ1 DNA sequence.
KW expression cassette; inducible promoter; alcl; alcR; alda; regulator;
KW alcohol dehydrogenase; herbicide resistance gene; glycosidase; GOX; EPS5S;
KW N-phosphonethyl-glycine; glyphosate oxidase; GOX; EPS5S;

RESULT 6
12 AGGAGTCAATTCAATTGGATGGACA - CGTGTGCAATTCTCACAAATTACACA 69
Db 274 AGGAGTCAATTCAATTGGAGGAGCACGATCTTCGATCTTCACAAATTACACA 333
Oy 70 ACACAAACAAACAAACATTACATTACATTAACTTACATTACA 115
Db 334 ACAACAAACAAACAAACAT-TACAAATTACATTACATTACA 377

KW 5-enol-pyruvylshikimate 3-phosphate CP4; chloroplast transit sequence 1;
 XX Arabidopsis; RUBISCO; CPT1; dicotyledon; monocotyledon; crop; ss.
 OS Synthetic.
 OS Synthetic.
 XX WO9706269-A1.
 XX PD 20-FEB-1997.
 XX PF 02-AUG-1996; 96WO-GB001883.
 XX PR 03-AUG-1995; 95GB-00015941.
 XX PA (ZENE) ZENECA LTD.
 XX PI Jebson I;
 XX DR WPI; 1997-154273/14.
 PR Chemically inducible cassette for expressing herbicide resistance gene in plants - and derived plants, partic. for resistance to glyphosate, avoids constitutive expression and minimises development of herbicide tolerant weeds.
 XX PS Example 3; Fig 8; 59pp; English.
 XX A chemically inducible plant gene expression cassette which comprises an inducible promoter linked to a gene (I) that confers resistance to a herbicide, is claimed. (I) imparts resistance to N-phosphonomethyl-glycine (glyphosate) or its salts, but may also be a gene for resistance to, e.g. chloroacetanilides, glufosinate, sulphonyl ureas, imidazolones, etc. The inducible promoter (e.g. alcR, alcA, aldB or other alcR-induced gene promoter) is operatively linked to an alcR regulator sequence. Induced expression of (I) avoids the risk that constitutive expression may interfere with plant development; allows volunteer plants to be controlled by herbicide applied without inducer and minimises the chance that herbicide-resistant strains of weeds will arise. The expression cassette is strictly controlled and suitable for general use in plants (both mono- and dicotyledons). The alcA/alcR gene switch was exemplified with genes conferring resistance to glyphosate. The switch was used to drive inducible expression of Glyphosate oxidase (GOX) in plants. Switchable GOX was expressed alone or in conjunction with constitutive expression of 5-enol-pyruvylshikimate 3'-phosphate (EPSPS) CP4. Constructs were optimised for expression in mono- and dicotyledonous crop species. The present sequence is that of plasmid PMJB1, based on pIBT21 containing the CaMV 35 promoter with duplicated enhancer linked to the tobacco mosaic virus translational enhancer sequence replacing the tobacco etch virus 5' non-translated leader, and terminated with the nopaline synthase poly(A) signal (nos). PMJB1 was used in construction of dicot vectors
 XX Sequence 1138 BP; 370 A; 253 C; 234 G; 281 T; 0 U; 0 Other;
 SQ Query Match 46.5%; Score 60.4; DB 2; Length 1138;
 Best Local Similarity 85.8%; Pred. No. 3.3e-05;
 Matches 91; Conservative 0; Mismatches 11; Indels 4; Gaps 2;

Qy 12 AGGAGTCTATTCTATTCATTGGAAATGGACAGCTGTGTTGTTGTTGATTTCACATTCACACAC 71
 Db 743 AGGAAGTCATTCTATTCATTGGAGG-ACCTCGAGTATTACTATTAGATTACATC 117
 Qy 72 AACAAACAAACAAACAAACATTAACATTACATTACATTACAC 800
 Db 801 AACRAACAAACAAACAT-TACATTACTATTACATTACAC 844
 RESULT 9
 AX78852 Standard; cDNA; 1138 BP.
 ID AX78852;
 XX AC AX78852;
 XX DT 07-SEP-1999 (first entry)
 AC AAT59713;
 XX DE Plasmid PMJB1 EcoRI-HindIII cDNA fragment.

DT 17-OCT-2003 (revised)
 XX 12-MAY-1997 (first entry)
 DE Plasmid pMJB1 used for use in delta-endotoxin expression.
 XX Gene expression cassette; promoter; alcR regulator; insecticide;
 KW CryIA(c); CryV; crystal protein; delta-endotoxin; Bacillus thuringiensis;
 KW Lepidoptera; insect resistance; transgenic plant; pMJB1; crop protection;
 KW biological control; alcR; gene switch; ds.
 XX OS Brassica oleracea var. botrytis; mosaic virus.
 OS Nicotiana tabacum; mosaic virus.
 OS Agrobacterium tumefaciens.
 OS Chimeric.
 XX FH Location/Qualifiers
 PI Key
 PN TATA_signal
 XX FT 737 . 743
 XX FT /*tag= C
 XX PN WO9706268-A2.
 XX PD 20-FEB-1997.
 XX PR 29-JUL-1996; 96WO-GB001846.
 XX PR 08-AUG-1995; 95GB-00016241.
 XX PA (ZENE) ZENECA LTD.
 XX PI Jebson I, Paine JAM;
 XX DR WPI; 1997-154272/14.
 XX PR Chemically inducible expression cassette - contains inducible promoter activated by alcR regulator in presence of alcohol or ketone inducer, used for insecticide production in plants.
 XX PS Example 4; Fig 11; 52pp; English.
 XX CC Plasmid pMJB1 (AAT59713) is based on pIBT21 and contains the cauliflower mosaic virus 35S promoter with duplicated enhancer linked to the tobacco mosaic virus (TMV) translational enhancer sequence and a nopaline-symthase 3' gene region. Incorporation of an endotoxin CryV (see also AAT59702) or CryIA(c) (see also AAT59712) gene into the plasmid downstream of the TMV enhancer provides vectors for the constitutive or inducible expression of CryV or CryIA(c)A. Transformation of plants allows the generation of transgenic plants resistant to insect pests. (Updated on 17-OCT-2003 to standardise OS field)
 XX SQ Sequence 1138 BP; 369 A; 254 C; 281 T; 0 U; 0 Other;
 SQ Query Match 46.5%; Score 60.4; DB 2; Length 1138;
 Best Local Similarity 85.8%; Pred. No. 3.3e-05;
 Matches 91; Conservative 0; Mismatches 11; Indels 4; Gaps 2;

Qy 12 AGGAAGTCATTCTATTCATTGGAAATGGACAGCTGTGTTGTTGTTGATTTCACATTCACACAC 71
 Db 743 AGGAAGTCATTCTATTCATTGGAGG-ACCTCGAGTATTACTATTAGATTACATC 117
 Qy 72 AACAAACAAACAAACAAACATTAACATTACATTACAC 800
 Db 801 AACRAACAAACAAACAT-TACATTACTATTACATTACAC 844
 RESULT 9
 AX78852 Standard; cDNA; 1138 BP.
 ID AX78852;
 XX AC AX78852;
 XX DT 07-SEP-1999 (first entry)
 AC AAT59713;
 XX DE Plasmid PMJB1 EcoRI-HindIII cDNA fragment.

XX	Paraquat; antibody; light chain; herbicide; resistant; crop plant; weed control; tolerant; diquat; photosynthesis inhibitor; photosystem I; free radical; lipid peroxidation; electron transport; photosystem II; vacuole; cell surface; cytotoxic; sensitive; ss.	XX	Cauliflower mosaic virus 3SS RNA; 3SS promoter; transgenic plant; OPR; KW
KW		KW	chlorotic streak cauliflower mosaic virus full length transcript promoter; PCISV; 12'-oxophytodienoic acid-100, 11-reductase gene promoter; SAG12; transgenic; CaMV; chimeric; ds.
KW		KW	
XX	Synthetic.	XX	Peanut; chloritic streak cauliflower virus.
OS		OS	Cauliflower mosaic virus.
XX		XX	Chimeric.
PN	WO932650-A1.	PN	
XX		XX	US6388170-B1.
PD	01-JUL-1999.	PD	
XX		XX	14-MAY-2002.
PF	15-DEC-1998; 9BMO-GB003760.	PF	
XX		XX	07-APR-2000; 2000US-00545244.
PR	19-DEC-1997; 976B-00026955.	XX	
XX		XX	07-APR-2000; 2000US-00545244.
PA	(ZENECA LTD.	PA	
XX		XX	(KENT) UNIV KENTUCKY RES FOUND.
P1	Holt DC, Jones PG;	P1	
XX		XX	Gan S, Xie M, He Y;
DR	WPI; 1999-405173/34.	XX	
XX		XX	WPI; 2002-442888/47.
PT	Herbicide binding proteins and related polynucleotides.	PT	
XX		XX	Bi-directional promoter with common promoter linked in opposite orientation to minimal promoter, useful to direct expression of polymucleotide which confers agronomically significant trait to transgenic plant.
PS	Disclosure; Page 46-47; 60pp; English.	PS	
XX		XX	Bi-directional promoter with common promoter having a minimal promoter operably linked in a bi-directional promoter comprising
CC		CC	CC a bi-directional promoter having a minimal promoter operably linked in opposite orientation 5' to a common promoter, where each promoter is either cauliflower mosaic virus (CaMV) 3SS RNA promoter (35S), peanut
CC		CC	CC chlorotic streak cauliflower virus full length transcript promoter (PCISV).
CC		CC	Arabidopsis 12'-oxophytodienoic acid-100, 11-reductase gene promoter (OPR)
CC		CC	or SAG12, and where each end of the bi-directional promoter is operably linked to a polynucleotide encoding a polypeptide. The construct is used
CC		CC	CC to produce a transgenic plant which has an agronomically-significant trait. The present sequence is pGL482 construct comprising 35S-PCISV
CC		CC	CC promoter DNA. (Updated on 29-AUG-2003 to standardize OS field)
CC		XX	Sequence 157 BP; 57 A; 38 C; 22 G; 40 T; 0 U; 0 Other;
CC		XX	Query Match 45.8%; Score 59.6; DB 6; Length 157;
CC		XX	Best Local Similarity 87.3%; Pred. No. 4.e-05;
CC		XX	Mismatches 89; Conservative 0; Mismatches 9; Indels 4; Gaps 2;
CC		Qy	12 AGGAAGTCATTCAATTGGATGGACA--CGTGTGTCATTCTCACAAATTACCAACA 69
CC		Db	41 AGGAAGTCATTCAATTGGAGGAGCAGGCTCTGAGACCTTACACATTACACA 100
CC		Qy	70 ACAACAAACAACAAACATTATAACATTACTATTACAAAT 111
CC		Db	101 ACAACAAACAACAAACAT-TACATTACTATTACAAAT 140
CC			RESULT 1
CC			ABL58982
CC			ID ABL58982 standard; DNA; 2200 BP.
CC			XX
CC			XX
CC			22-JUL-2002 (first entry)
AC	AAD44423 standard; DNA; 157 BP.	XX	
AC	AAD44423;	XX	
XX		XX	HPV16-L1 encoding DNA 2.
DT	29-AUG-2003 (revised)	XX	Human papilloma virus; HPV; HPV16-L1; HPV16-L2; vaccine; cervical cancer;
DT	13-DEC-2002 (first entry)	XX	HPV16-L1; cyostatic; virucide; gene; ds.
XX		XX	Human papillomavirus.
DE	pGL482 construct comprising 35S-PCISV promoter DNA.	OS	Synthetic.

DB CDNA sequence of the genome of tobacco mosaic virus-UI.
 XX Plant phenotype; gene trait; Nicotiana; Oryza sativa; Zea mays; Brassica;
 KW Gossypium; Triticum; Arabidopsis; Petunia; herbicide; transgenic plant;
 KW tobacco necrosis virus; TMV; tobacco mosaic virus; TMV; helper virus; ss.
 XX OS Tobacco mosaic virus.
 XX PN WO200063397-A2.
 XX PD 26-OCT-2000.
 XX PF 17-APR-2000; 2000WO-EP003521.
 XX PR 20-APR-1999; 99US-00284022.
 XX PA (AVET) AVENTIS CROPSCIENCE NV.
 XX PI Meulewaeter F, Cornelissen M, Jacobs J, Van Eldik G, Metzlaff M;
 XX DR WPI; 2000-687182/67.
 XX PT Identifying and isolating genes involved in determining the trait or
 PR phenotype of plant species, by infecting plants with gene silencing
 PR constructs targeted to the gene, and identifying plants with altered
 PR traits.

Example 1: Page 50-52: 64pp; English.

XX The specification describes a method for isolating genes that determine a trait or phenotype of a plant species. The method comprises identifying a set of nucleic acids of genes correlated with the trait, creating a library of gene silencing constructs in a viral RNA vector, targeting the gene silencing constructs to the nucleic acid set, infecting a collection of individual plants with these, identifying plants with altered traits or phenotype, and isolating genes of the invention. The method is useful for isolating genes involved in the determination of trait or a phenotype of a plant such as Nicotiana, Oryza sativa, Zea mays, Brassica, Gossypium, Triticum, Arabidopsis or Petunia. The method is also useful for validating the expression of selected nucleic acid sequences and for correlating the function of a nucleic acid sequence whose expression is correlated with the presence or absence of a specific trait in plants, but with otherwise unknown function. The method is also useful for developing agronomically useful products such as herbicides or transgenic plants. The present sequence represents the cDNA sequence of the genome of tobacco mosaic virus (TMV). The sequence encodes a movement protein, and was used to construct a plasmid vector for the synthesis of an infective hybrid tobacco mosaic virus (TMV)/ tobacco necrosis virus (TNV) helper virus RNA. This helper virus is used in the method of the invention.

XX Sequence 6395 BP; 1862 A; 1223 C; 1545 G; 1765 T; 0 U; 0 Other;

Query Match 40.0%; Score 52; DB 3;
 Best Local Similarity 85.4%; Pred. No. 0.0037;
 Matches 70; Conservative 0; Mismatches 10; Indels 2; Gaps 1;
 QY 46 GTCAATTCTAACATTACAAACAAACAAACAAACATATAAACTTACTT 105
 Db 1 GTATTTTAACATTACAAACAAACAAACAAACATATAAACTTACTT 58

QY 106 TACAAATTACATCTAGATTAACA 127
 Db 59 TACAAATTACATCTAGGCTATACCA 80

Search completed: July 3, 2005, 02:38:07
 Job time : 202.647 secs

This Page Blank (use back of page)

Copyright GenCore version 5.1.6
(c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 02:18:08 ; Search time 58.5294 Seconds
(without alignments)

3634.343 Million cell updates/sec

Title: US-10-814-858a-2

Perfect score: 130

Sequence: 1 tcactatataaggaaatcc.....ttacatcttagataaaatcg 130

Scoring table: IDENTITY NUC Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0\$ Maximum Match 100\$

Listing first 45 summaries

Database : Issued_Patents_NA:\$

1: /cgn2_6/podata/1/ina/5A_COMB.seq:\$
2: /cgn2_6/podata/1/ina/5B_COMB.seq:\$
3: /cgn2_6/podata/1/ina/6A_COMB.seq:\$
4: /cgn2_6/podata/1/ina/6B_COMB.seq:\$
5: /cgn2_6/podata/1/ina/POTUS_COMB.seq:\$
6: /cgn2_6/podata/1/ina/backfile1.seq:\$

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	130	100.0	452	4	US-09-263-692A-1
2	89	68.5	89	4	US-09-263-692A-13
3	60.4	46.5	1138	3	US-09-011-151-8
c	4	60.4	46.5	1138	3 US-09-011-151-9
5	59.6	45.8	157	3	US-09-545-244A-9
6	56.2	43.2	5033	1	US-08-038-168A-1
7	52	40.0	6395	2	US-08-687-59-2
8	52	40.0	6395	3	US-09-259-741-1
9	52	40.0	6395	3	US-09-037-151-1
10	52	40.0	6395	3	US-09-466-422-1
11	52	40.0	6395	4	US-09-401-415-2
12	52	40.0	6395	4	US-09-962-127-1
13	52	40.0	6425	3	US-09-259-741-3
14	52	40.0	6425	3	US-09-037-751-3
15	52	40.0	6425	3	US-09-466-422-3
16	52	40.0	6425	4	US-09-962-127-3
17	52	40.0	6439	3	US-09-259-741-2
18	52	40.0	6439	3	US-09-037-151-2
19	52	40.0	6439	3	US-09-466-422-2
20	52	40.0	6439	4	US-09-962-527-2
21	52	40.0	6446	3	US-09-259-741-5
22	52	40.0	6446	3	US-09-037-151-5
23	52	40.0	6446	3	US-09-466-422-5
24	52	40.0	6446	4	US-09-962-127-5
25	52	40.0	6475	3	US-09-259-741-4
26	52	40.0	6475	3	US-09-037-751-4
27	52	40.0	6475	3	US-09-466-422-4

SEQUENCE ALIGNMENTS

RESULT 1
US-09-263-692A-1

; Sequence 1, Application US/09263692A
; Patent No. 6639065

; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A chemically synthesized artificial promoter for high level expr
; FILE REFERENCE: Q52511
; CURRENT APPLICATION NUMBER: US/09/263, 692A
; CURRENT FILING DATE: 1999-03-05
; PRIORITY APPLICATION NUMBER: 3322/Del/98
; PRIORITY FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 452
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic DNA promoter sequence

US-09-263-692A-1

Query Match 100.0\$ Score 130; DB 4; Length 452;
Best Local Similarity 100.0\$ Pred. No. 1.1e-2;
Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCACTATATAAGGAGTCATTGAGTGGATTTGAGCTGTGATTTCTCAACAA 60
Db 313 TCACATATATAAGGAGTCATTGAGTGGATTTGAGCTGTGATTTCTCAACAA 372
Qy 61 TTACCAACACAAACAAACACAAACATTATAATTACTATTACATTACATCTAG 120
Db 373 TTACCAACACAAACAAACACAAACATTATAATTACTATTACATTACATCTAG 432
Qy 121 ATACAATAG 130
Db 433 ATACAATAG 442

RESULT 2
US-09-263-692A-13

; Sequence 13, Application US/09263692A
; Patent No. 6639065

; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A chemically synthesized artificial promoter for high level expr
; FILE REFERENCE: Q52511
; CURRENT APPLICATION NUMBER: US/09/263, 692A

CURRENT FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: 33222/De1/98
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 13
; LENGTH: 89
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: 5' untranslated leader region

US-09-263-692A-13

Query Match 68.5%; Score 89; DB 4; Length 89;
Best Local Similarity 100.0%; Pred. No. 3e-14; Indels 0; Gaps 0;
Matches 89; Conservative 0; Mismatches 0; Gaps 0;

Qy 39 ACGGTTGATTCATTCAAACTTACCAACAAACAAACATTATACAT 98
Db 1 ACGGTTGATTCATTCAAACTTACCAACAAACAAACATTATACAT 60

Qy 99 TACTATTACAAATTACATCTAGATAACA 127
Db 61 TACTATTACAAATTACATCTAGATAACA 89

RESULT 3
US-09-011-151-8
; Sequence 8, Application US/09011151
; Patent No. 6380163
; GENERAL INFORMATION:
; APPLICANT: Jepson, Ian
; TITLE OF INVENTION: DNA Constructs
; FILE REFERENCE: PPD 50059/UST
; CURRENT APPLICATION NUMBER: US/09/011,151
; CURRENT FILING DATE: 1998-01-29
; PRIOR APPLICATION NUMBER: PCT/GB96/01883
; PRIOR FILING DATE: 1996-08-02
; PRIOR FILING DATE: 1995-08-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 8
; LENGTH: 1138
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Plasmid pMJB1

US-09-011-151-8

Query Match 46.5%; Score 60.4; DB 3; Length 1138;
Best Local Similarity 83.8%; Pred. No. 1e-06; Indels 4; Gaps 2;
Matches 91; Conservative 0; Mismatches 11; Gaps 2;

Qy 12 AGGAAGTCATTCAATTGGAAATGGATGACAGCTGTGTCATTCTCAAGATTACAAACAC 71
Db 743 AGGAAGTCATTCAATTGGAAATGGACAGCTGTGTCATTCTCAACATTACAAACAC 71

Qy 72 AACAAACAAACAAACATTATACTTACATTACATC 117
Db 801 AACAAACAAACAAACAT-TACATTACTTACATTACAC 844

RESULT 4
US-09-011-151-9/c
; Sequence 9, Application US/09011151
; Patent No. 6380163
; GENERAL INFORMATION:
; APPLICANT: Jepson, Ian
; TITLE OF INVENTION: DNA Constructs
; FILE REFERENCE: PPD 50059/UST
; CURRENT APPLICATION NUMBER: US/09/011,151
; CURRENT FILING DATE: 1998-01-29

CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS, Microsoft Windows
SOFTWARE: Microsoft Windows No. 5569828epad
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/038,768A
FILING DATE: 19930324
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Roth, Michael J.
REGISTRATION NUMBER: 29,342
REFERENCE DOCKET NUMBER: 0235 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 245-3594
TELEFAX: (515) 245-3634
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5033 base pairs
TYPE: nucleotide
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: synthetic DNA
DESCRIPTION: transformation plasmid pPHI1406
HYPOOTHECAL: No
-08-038-768A-1

Query Match	Score	DB 1:	Length
Best Local Similarity	43.2%	DB 1:	Length 5033;
Local Similarity	83.2%	Pred. No.	1.6e-05;
Conservative Matches	89	N mismatches	13;
		Indels	5
1165 AGGAGTTCATTCTATGGAAATGGAC--ACGTGTTGTCATTCTCAACAA			
1165 AGGAGTTCATTCTATGGAGAGGAGGAGCTGCAGCTTATTTCACAAAC			

69 AACAAACAAACAAACAAACAAACATACTACATTACATTACA 115
 1225 AAGAACAAACAAACAAACAAACAT--TACATTACTATTACATTACA 1269

RESULT 7

-08-887-559-2 Sequence 2, Application US/08687559

Patient No. 5955647 GENERAL INFORMATION:

APPLICANT: Beachy, Roger N.
TITLE OF INVENTION: METHOD FOR USING TOBACCO MOSAIC VIRUS TO OVERPRODUCE PEPTIDES AND PROTEINS
NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,559
FILING DATE: No. 5955647ember 18, 1996
CLASSIFICATION: 424

```

PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01467
FILING DATE: 03-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Learni, June M.
REGISTRATION NUMBER: 31,238
REFERENCE/DOCKET NUMBER: 07302/011001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-678-5070
TELEFAX: 619-678-5099
SEQUENCE CHARACTERISTICS:
LENGTH: 6395 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
IMMEDIATE SOURCE:
CLONE: TMV
US-08-687-559-2

Query Match Score 52; DB 2; Length 6395;
Best Local Similarity 40.0%; Pred. No. 0.00019;
Matches 70; Conservative 0; Mismatches 10; Indels 2; Gaps 1;
Qy      46 GTCAATTCGAACTTACCAACAACTAACAAACAAACATTAACATTACTAT 105
Db      1 GTATTTTACAACTTACCAACAAACAAACAAACAT--TACAACTTACTAT 58
Qy      106 TACAAATTACATCTAGATAAACAA 127
Db      59 TACATTACATGGCATACCA 80

```

RESULT 8
US-09-259-741-1
Sequence 1, Application US/09259741
Patent No. 603895
GENERAL INFORMATION:
APPLICANT: GARGER, STEPHEN
APPLICANT: HOLTZ, R. BARRY
APPLICANT: MCCULLOCH, MICHAEL
APPLICANT: TURPEN, THOMAS
TITLE OF INVENTION: A PROCESS FOR ISOLATING AND PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES FROM PLANT SOURCES
TITLE OF INVENTION: NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howrey & Simon
STREET: 1299 Pennsylvania Avenue N.W.
CITY: Washington
STATE: DC

ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEO for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/259,741
FILING DATE: February 25, 1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/037,751
FILING DATE: March 10, 1998
ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P
REGISTRATION NUMBER: 25,277
REFERENCE/DOCKET NUMBER: 00801-0140.US01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-463-8100
TELEFAX: 650-463-0000

TELEX:
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 6395 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: Genomic RNA
 US-09-255-741-1

Query Match 40.0%; Score 52; DB 3; Length 6395;
 Best Local Similarity 64.6%; Pred. No. 0.00019;
 Matches 53; Conservative 17; Mismatches 10; Indels 2; Gaps 1;

Qy 46 GTCAATTCTCACAAATPACCAAACAAACAAACATTAACATTACTATT 105
 Db 1 GUAUUUAACAUACAUACAAACAAACAAACAAACAAACAAACAU--UACAUUACAUU 58

Qy 106 TACAAATCATCTGATAAAC 127
 Db 59 UACAUUACAUGGCAUACAA 80

RESULT 9
 US-09-037-751-1
 Sequence 1, Application US/09037751
 Patent No. 603745
 GENERAL INFORMATION:
 APPLICANT: GARGER, STEPHEN
 APPLICANT: HOLTZ, R. BARRY
 APPLICANT: MCCULLOCH, MICHAEL
 APPLICANT: TURPEN, THOMAS
 TITLE OF INVENTION: A PROCESS FOR ISOLATING AND PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES
 NUMBER OF SEQUENCES: 5
 TITLE OF INVENTION: A PROCESS FOR ISOLATING AND PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES
 NUMBER OF SEQUENCES: 5
 TITLE OF INVENTION: A PROCESS FOR ISOLATING AND PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Howrey & Simon
 STREET: 1299 Pennsylvania Avenue N.W.
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20004
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/466,422
 FILING DATE: 17-Dec-1999
 CLASSIFICATION: <Unknown>
 PRIORITY DATA:
 APPLICATION NUMBER: 09/037,751
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Halluin, Albert P
 REGISTRATION NUMBER: 25,277
 REFERENCE/DOCKET NUMBER: 00801.0140.999
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-463-8109
 TELEFAX: 650-463-8400
 TELEX: <Unknown>
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 6395 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: Genomic RNA
 US-09-037-751-1

Query Match 40.0%; Score 52; DB 3; Length 6395;
 Best Local Similarity 64.6%; Pred. No. 0.00019;
 Matches 53; Conservative 17; Mismatches 10; Indels 2; Gaps 1;

Qy 46 GTCAATTCTCACAAATPACCAAACAAACAAACATTAACATTACTATT 105
 Db 1 GUAUUUAACAUACAUACAAACAAACAAACAAACAU--UACAUUACAUU 58

Qy 106 TACAAATCATCTGATAAAC 127
 Db 59 UACAUUACAUGGCAUACAA 80

RESULT 11
US-03-401-415-2
Sequence 2, Application US/09401415
Patent No. 6503732
GENERAL INFORMATION:
APPLICANT: THE SCRIPPS RESEARCH INSTITUTE
TITLE OF INVENTION: METHOD FOR USING TOBACCO MOSAIC VIRUS TO
OVERPRODUCE PEPTIDES AND PROTEINS
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Lubitz
STREET: 1880 Century Park East, Suite 500
CITY: Los Angeles
STATE: California
ZIP: 90067
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/401,415
FILING DATE: 21-Sep-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01467
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Boblich, June M.
REGISTRATION NUMBER: 31,238
REFERENCE/DOCKET NUMBER: FD-4074
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 6395 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic RNA
IMMEDIATE SOURCE:
CLONE: TMV
FEATURE:
NAME/KEY: CDS
LOCATION: 1..6395
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-401-415-2

Query Match 40.0%; Score 52; Length 6395;
Best Local Similarity 85.4%; Pred. No. 0.00019;
Matches 70; Conservative 0; Mismatches 10; Indels 2; Gaps 1;

Qy 46 GTCAATTCTCAACATTAACAAACACAAACAAACATTATAACATTACTATT 105
Db 1 GTATTTTACAACTTACAAACAAACAAACAAACAAACAT--TACAATTACTATT 58

RESULT 12
US-09-962-527-1
Sequence 1, Application US/0962527
Patent No. 6740740
GENERAL INFORMATION:
APPLICANT: GARGER, STEPHEN
HOLTZ, R. BARRY
MCCULLOCH, MICHAEL
APPLICANT: TURPEN, THOMAS
APPLICANT: MCCULLOCH, MICHAEL
APPLICANT: GARGER, STEPHEN
HOLTZ, R. BARRY
APPLICANT: MCCULLOCH, MICHAEL
APPLICANT: TURPEN, THOMAS
APPLICANT: MCCULLOCH, MICHAEL
APPLICANT: GARGER, STEPHEN
HOLTZ, R. BARRY
APPLICANT: MCCULLOCH, MICHAEL

RESULT 13
US-09-962-741-3
Sequence 3, Application US/09259741
Patent No. 6033895
GENERAL INFORMATION:
APPLICANT: GARGER, STEPHEN
HOLTZ, R. BARRY
APPLICANT: MCCULLOCH, MICHAEL
APPLICANT: TURPEN, THOMAS
APPLICANT: MCCULLOCH, MICHAEL
APPLICANT: GARGER, STEPHEN
HOLTZ, R. BARRY
APPLICANT: MCCULLOCH, MICHAEL
APPLICANT: TURPEN, THOMAS
APPLICANT: MCCULLOCH, MICHAEL
APPLICANT: GARGER, STEPHEN
HOLTZ, R. BARRY
APPLICANT: MCCULLOCH, MICHAEL

TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES
FROM PLANT SOURCES
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howrey & Simon
STREET: 1299 Pennsylvania Avenue N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/962,527
FILING DATE: 24-Sep-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/037,751
FILING DATE: 10-March-1998
ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P.
REGISTRATION NUMBER: 25,277
REFERENCE/DOCKET NUMBER: 00801.0140.999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-463-8109
TELEFAX: 650-463-8400
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6395 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: Genomic RNA
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-962-527-1

Query Match 40.0%; Score 52; DB 4; Length 6395;
Best Local Similarity 64.6%; Pred. No. 0.00019;
Matches 53; Conservative 17; Mismatches 10; Indels 2; Gaps 1;

Qy 46 GTCAATTCTCAACATTAACAAACACAAACAAACATTACTATT 105
Db 1 GUAUUUTACAAACAAACAAACAAACAAACAAACAAACAUU 58

Query Match 40.0%; Score 52; Length 6395;
Best Local Similarity 85.4%; Pred. No. 0.00019;
Matches 70; Conservative 0; Mismatches 10; Indels 2; Gaps 1;

Qy 106 TACATTACATCTGATAAACAA 127
Db 59 UACAUUACAAUUGCAUCACA 80

COUNTRY: USA
ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ For Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/259,741
FILING DATE: February 25, 1999
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 09/037,751
FILING DATE: March 10, 1998
ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P
REGISTRATION NUMBER: 25,277
REFERENCE/DOCKET NUMBER: 00801.0140.999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-463-8109
TELEFAX: 650-463-8400
TELEX:
US-09-037-751-3

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 6425 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: Genomic RNA

Query Match 40.0%; Score 52; DB 3; Length 6425;
Best Local Similarity 64.6%; Pred. No. 0.0002;
Matches 53; Conservative 17; Mismatches 10; Indels 2; Gaps 1;

Qy 46 GTCAATTCTCACATACCAACACAAACAAACATAACATTACATT 105
Db 1 GUAUUUTAACAGAACAUACACAAACAAACACAU-UACAAUACAUU 58

Qy 106 TAGATTACATCTAGATAAACA 127
Db 59 UACAAUACAUAGGCAUACAA 80

RESULT 14
US-09-037-751-3
Sequence 3 Application US/09037751
Patent No. 603456

GENERAL INFORMATION:
APPLICANT: GARGER, STEPHEN
APPLICANT: HOLITZ, R. BARRY
APPLICANT: MCCULLOCH, MICHAEL
APPLICANT: TURPEN, THOMAS

TITLE OF INVENTION: A PROCESS FOR ISOLATING AND PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES
TITLE OF INVENTION: PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES
NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:
ADDRESSEE: Howrey & Simon
STREET: 1299 Pennsylvania Avenue N.W.
CITY: Washington
STATE: DC
COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ For Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/466,422

FILING DATE: 17-Dec-1999
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 09/037,751

FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P
REGISTRATION NUMBER: 25,277
REFERENCE/DOCKET NUMBER: 00801.0140.999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-463-8109
TELEFAX: 650-463-8400
TELEX: <Unknown>

RESULT 15
US-09-466-422-3
Sequence 3 Application US/09466422
Patent No. 6303779

GENERAL INFORMATION:
APPLICANT: GARGER, STEPHEN
APPLICANT: HOLITZ, R. BARRY
APPLICANT: MCCULLOCH, MICHAEL
APPLICANT: TURPEN, THOMAS

TITLE OF INVENTION: A PROCESS FOR ISOLATING AND PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES
FROM PLANT SOURCES

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:
ADDRESSEE: Howrey & Simon
STREET: 1299 Pennsylvania Avenue N.W.
CITY: Washington
STATE: DC
COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ For Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/466,422

FILING DATE: 17-Dec-1999
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 09/037,751

FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P
REGISTRATION NUMBER: 25,277
REFERENCE/DOCKET NUMBER: 00801.0140.999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-463-8109
TELEFAX: 650-463-8400
TELEX: <Unknown>

RESULT 16
US-09-037-751-3
Sequence 3 Application US/09037751
Patent No. 603456

GENERAL INFORMATION:
APPLICANT: GARGER, STEPHEN
APPLICANT: HOLITZ, R. BARRY
APPLICANT: MCCULLOCH, MICHAEL
APPLICANT: TURPEN, THOMAS

TITLE OF INVENTION: A PROCESS FOR ISOLATING AND PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES
TITLE OF INVENTION: PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES
TITLE OF INVENTION: FROM PLANT SOURCES

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:
ADDRESSEE: Howrey & Simon
STREET: 1299 Pennsylvania Avenue N.W.
CITY: Washington
STATE: DC
COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ For Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/466,422

FILING DATE: 17-Dec-1999
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 09/037,751

FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P
REGISTRATION NUMBER: 25,277
REFERENCE/DOCKET NUMBER: 00801.0140.999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-463-8109
TELEFAX: 650-463-8400
TELEX: <Unknown>

This Page Blank (use back)

Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 03:19:18 ; Search time 266.765 Seconds
(without alignments)

3056.566 Million cell updates/sec

Title: US-10-814-858a-2

Perfect score: 130

Sequence: 1 tcacatataataggaaatcc.....ttacatctagataaaactg 130

Scoring table: IDENTITY_NUC Gapop 10.0 , Gapext 1.0

Searched: 6133374 seqs, 3136092125 residues

Total number of hits satisfying chosen parameters: 12626748

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA:*

1: /cgn2_6/ptodata/1/pubpna/us07_pubcomb.seq:*

2: /cgn2_6/ptodata/1/pubpna/pct_new_pub.seq:*

3: /cgn2_6/ptodata/1/pubpna/us06_pubcomb.seq:*

4: /cgn2_6/ptodata/1/pubpna/us07_new_pub.seq:*

5: /cgn2_6/ptodata/1/pubpna/pctus7_new_pub.seq:*

6: /cgn2_6/ptodata/1/pubpna/us08_pubcomb.seq:*

7: /cgn2_6/ptodata/1/pubpna/us08_new_pub.seq:*

8: /cgn2_6/ptodata/1/pubpna/us08_pubcomb.seq:*

9: /cgn2_6/ptodata/1/pubpna/us09a_pubcomb.seq:*

10: /cgn2_6/ptodata/1/pubpna/us09b_pubcomb.seq:*

11: /cgn2_6/ptodata/1/pubpna/us09c_pubcomb.seq:*

12: /cgn2_6/ptodata/1/pubpna/us09_new_pub.seq:*

13: /cgn2_6/ptodata/1/pubpna/us10a_pubcomb.seq:*

14: /cgn2_6/ptodata/1/pubpna/us10b_pubcomb.seq:*

15: /cgn2_6/ptodata/1/pubpna/us10c_pubcomb.seq:*

16: /cgn2_6/ptodata/1/pubpna/us10d_pubcomb.seq:*

17: /cgn2_6/ptodata/1/pubpna/us10e_pubcomb.seq:*

18: /cgn2_6/ptodata/1/pubpna/us10f_pubcomb.seq:*

19: /cgn2_6/ptodata/1/pubpna/us10g_pubcomb.seq:*

20: /cgn2_6/ptodata/1/pubpna/us10h_pubcomb.seq:*

21: /cgn2_6/ptodata/1/pubpna/us10i_pubcomb.seq:*

22: /cgn2_6/ptodata/1/pubpna/us10_new_pub.seq:*

23: /cgn2_6/ptodata/1/pubpna/us11a_pubcomb.seq:*

24: /cgn2_6/ptodata/1/pubpna/us11_new_pub.seq:*

25: /cgn2_6/ptodata/1/pubpna/us60_new_pub.seq:*

26: /cgn2_6/ptodata/1/pubpna/us60_pubcomb.seq:*

RESULT 1 US-10-784-295-13

; Sequence 13, Application US/10784295

; Publication No. US2005003221A1

; GENERAL INFORMATION:

; APPLICANT: Shaaltiel, Joseph

; TITLE OF INVENTION: CELL/TISSUE CULTURING DEVICE, SYSTEM AND METHOD

; FILE REFERENCE: 27557

; CURRENT APPLICATION NUMBER: US/10/784,295

; CURRENT FILING DATE: 2004-02-24

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 13

; LENGTH: 2186

; TYPE: DNA

; ORGANISM: Artificial sequence

; FEATURE:

; OTHER INFORMATION: Nucleic acid encoding recombinant GCD fused to signal peptides

; NAME/KEY: misc feature

; LOCATION: (2181).(2181)

; OTHER INFORMATION: n is a, c, g, or t

Qy 12 AGGAAGTCATTCACTTGGATGAGACA-CGTGTTGTCATTCTCACAAATTACCAACA 69

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Match	Length	DB	ID	Description
1	63.6	48.9	2186	21	US-10-784-295-13	Sequence 13, Appl
2	54.8	42.2	368	9	US-09-331-631A-29	Sequence 29, Appl
3	54.8	42.2	368	16	US-10-147-095-29	Sequence 29, Appl
4	52	40.0	6395	10	US-09-962-527-1	Sequence 1, Appl
5	52	40.0	6395	17	US-10-338-592-2	Sequence 2, Appl
6	52	40.0	6395	19	US-10-822-029-1	Sequence 1, Appl
7	52	40.0	6425	10	US-09-962-527-3	Sequence 3, Appl

Db 312 AGGAGTTCATTCAATTGGAGGACAGGCTTCTTGAGATCCCTAACAAATTACCAACA 371
 Qy 70 ACACAAACACAAACAAACATAACATACTATTAACTTAACTTAACTACA 115
 Db 372 ACAACAAACAAACAAACATACTTAACTTAACTTAACTACA 415

RESULT 2
 US-09-331-631A-29 Sequence 29, Application US/09331631A
 Patent No. US20020163392A1
 GENERAL INFORMATION:
 APPLICANT: Manners, John M.
 APPLICANT: Marcus, John Paul
 APPLICANT: Goult, Kenneth C.
 APPLICANT: Green, Jodie L.
 TITLE OF INVENTION: ANTIMICROBIAL PROTEINS
 FILE REFERENCE: CULLN3.001APC
 CURRENT APPLICATION NUMBER: US/09/331,631A
 CURRENT FILING DATE: 1999-06-21
 PRIOR APPLICATION NUMBER: PCT/AU97/00874
 PRIOR FILING DATE: 1997-12-22
 PRIOR APPLICATION NUMBER: AU PO 4275
 PRIOR FILING DATE: 1996-12-20
 NUMBER OF SEQ ID NOS: 40
 SOFTWARE: FastSEQ for Windows Version 3.0
 SEQ ID NO: 29 LENGTH: 368
 TYPE: DNA ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: A synthetic nucleotide sequence which can be used for the expression and secretion of MiAMP2c, containing the leader sequence from SEQ ID NO:11
 OTHER INFORMATION: for the expression and secretion of MiAMP2c, containing the leader sequence from SEQ ID NO:11
 OTHER INFORMATION: containing the leader sequence from SEQ ID NO:11
 OTHER INFORMATION: and SEQ ID NO:5.

Query Match 42.2%; Score 54.8; DB 16; Length 368;
 Best Local Similarity 76.9%; Pred. No. 0.00064;
 Matches 80; Conservative 0; Mismatches 22; Indels 2; Gaps 1;
 NAME/KEY: CDS
 LOCATION: (103) . . . (333)
 US-10-147-095-29

Query Match 42.2%; Score 54.8; DB 16; Length 368;
 Best Local Similarity 76.9%; Pred. No. 0.00064;
 Matches 80; Conservative 0; Mismatches 22; Indels 2; Gaps 1;
 NAME/KEY: CDS
 LOCATION: (103) . . . (333)

Query Match 42.2%; Score 54.8; DB 9; Length 368;
 Best Local Similarity 76.9%; Pred. No. 0.00064;
 Matches 80; Conservative 0; Mismatches 22; Indels 2; Gaps 1;
 NAME/KEY: CDS
 LOCATION: (103) . . . (333)

Query Match 42.2%; Score 54.8; DB 16; Length 368;
 Best Local Similarity 76.9%; Pred. No. 0.00064;
 Matches 80; Conservative 0; Mismatches 22; Indels 2; Gaps 1;
 NAME/KEY: CDS
 LOCATION: (103) . . . (333)

RESULT 3
 US-10-147-095-29 Sequence 29, Application US/10147095
 Publication No. US20030171274A1
 GENERAL INFORMATION:
 APPLICANT: Manners, John M.
 APPLICANT: Marcus, John Paul
 APPLICANT: Goult, Kenneth C.
 APPLICANT: Green, Jodie L.
 TITLE OF INVENTION: ANTIMICROBIAL PROTEINS
 FILE REFERENCE: CULLN3.001APC
 CURRENT APPLICATION NUMBER: US/10/147,095
 CURRENT FILING DATE: 2002-05-15
 PRIOR APPLICATION NUMBER: US/09/331,631A
 PRIOR FILING DATE: 1999-06-21
 PRIOR APPLICATION NUMBER: PCT/AU97/00874
 PRIOR FILING DATE: 1997-12-22
 PRIOR APPLICATION NUMBER: AU PO 4275
 NUMBER OF SEQ ID NOS: 40
 SOFTWARE: FastSEQ for Windows Version 3.0

Query Match 42.2%; Score 54.8; DB 16; Length 368;
 Best Local Similarity 76.9%; Pred. No. 0.00064;
 Matches 80; Conservative 0; Mismatches 22; Indels 2; Gaps 1;
 NAME/KEY: CDS
 LOCATION: (103) . . . (333)

RESULT 4
 US-09-362-527-1 Sequence 1, Application US/09962527
 Publication No. US20030049813A1
 GENERAL INFORMATION:
 APPLICANT: Garger, Stephen Holtz, R. Barry McCulloch, Michael Turpen, Thomas
 TITLE OF INVENTION: A PROCESS FOR ISOLATING AND PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES FROM PLANT SOURCES
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Howrey & Simon STREET: 1299 Pennsylvania Avenue N.W.
 CITY: Washington STATE: DC
 COUNTRY: USA ZIP: 20004
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS SOFTWARE: FastSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/962,527
 FILING DATE: 24-Sep-2001
 CLASSIFICATION: <Unknown>
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 09/037,751
 FILING DATE: 10-march-1998
 ATTORNEY/AGENT INFORMATION:
 NAME: Halluin, Albert P.
 REGISTRATION NUMBER: 25,277
 REFERENCE/DOCKET NUMBER: 25,277
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-463-8109
 TELEFAX: 650-463-8400
 TELEX: <Unknown>
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 6395 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: unknown

SEQUENCE CHARACTERISTICS:
 LENGTH: 6439 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: Genomic RNA
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-09-962-527-2

Query Match Score 52; DB 19; Length 6439;
 Best Local Similarity 64.6%; Pred. No. 0.007;
 Matches 53; Conservative 17; Mismatches 10; Indels 2; Gaps 1;

Qy 46 GTCAATTCTCAACATTAACAAACAAACAAACAAACAAACATTATAAACATTACTATT 105
 Db 1 GUADUUTUACAACAUACAUACAAACAAACAAACAAACAAACAAACAAACAAACAUU-UACAUUACAUU 58

Query Match Score 52; DB 10; Length 6439;
 Best Local Similarity 64.6%; Pred. No. 0.007;
 Matches 53; Conservative 17; Mismatches 10; Indels 2; Gaps 1;

Qy 46 GTCAATTCTCAACATTAACAAACAAACAAACAAACAAACATTATAAACATTACTATT 105
 Db 1 GUADUUTUACAACAUACAUACAAACAAACAAACAAACAAACAAACAAACAUU-UACAUUACAUU 58

SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-09-962-527-2

Query Match Score 52; DB 10; Length 6439;
 Best Local Similarity 64.6%; Pred. No. 0.007;
 Matches 53; Conservative 17; Mismatches 10; Indels 2; Gaps 1;

Qy 46 GTCAATTCTCAACATTAACAAACAAACAAACAAACAAACATTATAAACATTACTATT 105
 Db 1 GUADUUTUACAACAUACAUACAAACAAACAAACAAACAAACAAACAAACAUU-UACAUUACAUU 58

RESULT 11
 US-09-962-527-5

Sequence 5, Application US/099262527
 Publication No. US20030049813A1

GENERAL INFORMATION:
 ADDRESSER: Howrey & Simon
 STREET: 1299 Pennsylvania Avenue N.W.
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20004

TITLE OF INVENTION: A PROCESS FOR ISOLATING AND PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES FROM PLANT SOURCES

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Howrey & Simon
 STREET: 1299 Pennsylvania Avenue N.W.
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20004

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/962,527
 FILING DATE: 24-Sep-2001
 PRIORITY APPLICATION DATA:
 REFERENCE/DOCKET NUMBER: 09/037,751
 ATTORNEY/AGENT INFORMATION:
 NAME: Halluin, Albert P
 REGISTRATION NUMBER: 25,277
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-463-8109
 TELEFAX: 650-463-8400
 TELEX: <Unknown>
 INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:
 LENGTH: 6439 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: Genomic RNA
 SEQUENCE DESCRIPTION: SEQ ID NO: 5:
 US-09-962-527-5

Query Match Score 52; DB 10; Length 6446;
 Best Local Similarity 64.6%; Pred. No. 0.007;
 Matches 53; Conservative 17; Mismatches 10; Indels 2; Gaps 1;

Qy 46 GTCAATTCTCAACATTAACAAACAAACAAACAAACATTATAAACATTACTATT 105
 Db 1 GUADUUTUACAACAUACAUACAAACAAACAAACAAACAUU-UACAUUACAUU 58

SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-10-828-029-2

Query Match Score 52; DB 10; Length 6446;
 Best Local Similarity 64.6%; Pred. No. 0.007;
 Matches 53; Conservative 17; Mismatches 10; Indels 2; Gaps 1;

Qy 46 GTCAATTCTCAACATTAACAAACAAACAAACAAACATTATAAACATTACTATT 105
 Db 1 GUADUUTUACAACAUACAUACAAACAAACAAACAAACAUU-UACAUUACAUU 58

RESULT 12
 US-10-828-029-2

Sequence 2, Application US/10828029
 Publication No. US20040171813A1

GENERAL INFORMATION:
 APPLICANT: GARGER, STEPHEN
 MCCULLOCH, MICHAEL
 TURPEN, THOMAS

TITLE OF INVENTION: A PROCESS FOR ISOLATING AND PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES FROM PLANT SOURCES

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Howrey & Simon
 STREET: 1299 Pennsylvania Avenue N.W.
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20004

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/828,029
 FILING DATE: 20-Apr-2004
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US/09/962,527
 FILING DATE: 24-Sep-2001
 ATTORNEY/AGENT INFORMATION:
 NAME: Halluin, Albert P
 REGISTRATION NUMBER: 25,277
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-463-8109
 TELEFAX: 650-463-8400
 TELEX: <Unknown>
 INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:
 LENGTH: 6446 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: Genomic RNA
 SEQUENCE DESCRIPTION: SEQ ID NO: 5:
 US-10-828-029-2

Page 6

ADDRESSEE: Howrey & Simon
STREET: 1299 Pennsylvania Avenue N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ FOR Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/828,029
FILING DATE: 20-Apr-2004
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/962,527
FILING DATE: 24-Sep-2001
APPLICATION NUMBER: 09/037,751
FILING DATE: 10-march-1998

ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P
REGISTRATION NUMBER: 25,277
REFERENCE DOCKET NUMBER: 00801.0140.999

TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-453-8109
TELEFAX: 650-453-8400

TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 6475 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: unknown
MOLECULE TYPE: Genomic RNA
SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-10-828-029-4

Query Match 40.0%; Score 52; DB 14; Length 7926;
Best Local Similarity 85.4%; Pred. No. 0.0075;
Matches 70; Conservative 0; Mismatches 10; Indels 2; Gaps 1;

Qy 46 GTCAATTCTCAACATTACCAACAAACAAACATAATTACTATT 105
Db 1 GTATTTTACAACATTACCAACAAACAAACATAATTACTATT 105

106 TACATTACATCTAGATAAAC 127

Db 59 TACATTACATGCGATAGCA 80

Search completed: July 3, 2005, 07:33:16

Job time : 269.765 secs

US-10-828-029-4

Query Match 40.0%; Score 52; DB 19; Length 6475;

Best Local Similarity 64.6%; Pred. No. 0.007; Matches 53; Conservative 17; Mismatches 10; Indels 2; Gaps 1;

Qy 46 GTCAATTCTCAACATTACCAACAAACAAACATAATTACTATT 105
Db 1 GUADUUUAGACAUACACAAACAAACAAACAAACAU-UACAAUACAUU 58

106 TACATTACATCTAGATAAAC 127

Db 59 UACAAUACAAUGGCAACACA 80

RESULT 15

US-10-119-330-1
Sequence 1, Application US/10119330
Publication No. US20030073209A1

GENERAL INFORMATION:

APPLICANT: TURPEN, Thomas

APPLICANT: GARGER, Stephen

APPLICANT: MCCULLOCH, Michael

APPLICANT: CAMERON, Terri

APPLICANT: SAMNEK-POTTER, Michelle L.

APPLICANT: HOLTZ, R. Barry

TITLE OF INVENTION: METHOD FOR RECOVERING PROTEINS FROM THE

TITLE OF INVENTION: INTERSTITIAL FLUID OF PLANT TISSUES

FILE REFERENCE: 00801.0135.US00

CURRENT APPLICATION NUMBER: US/10/119,330

CURRENT FILING DATE: 2002-04-08

PRIOR APPLICATION NUMBER: US/09/132,989

PRIOR FILING DATE: 1998-08-11

NUMBER OF SEQ ID NOS: 1

SOFTWARE: FastSEQ for Windows Version 3.0

SEQ ID NO: 1

LENGTH: 7926

الطبقة العاملة (USA).